



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 145604

**TO: Minh-Tam Davis**  
**Location: REM-3A24&3C18**  
**Art Unit: 1642**  
**Thursday, March 03, 2005**

**Case Serial Number: 10/099791**

**From: Deirdre Arnold**  
**Location: Biotech-Chem Library**  
**REM 1A64**  
**Phone: 571-272-2532**

**Deirdre.Arnold@uspto.gov**

### Search Notes

Modified Request

*Please feel free to contact me if you have any questions or would like to amend the search.*

Thank you for using STIC services.

Regards,  
Deirdre Arnold



1

2

3

4

From: Davis, Minh-Tam  
Sent: Friday, February 18, 2005 2:27 PM  
To: STIC-Biotech/ChemLib  
Subject: Search request for 10/099791

CRFF

Please search in commercial database, issued patent files, PGPUB and interference:

1) SEQ ID NO:2

2) SEQ ID NO:4-11 (Fragments of SEQ ID NO:2), with and without size limitation of the sequences in database to the size of the corresponding sequences.

Thank you.

MINH TAM DAVIS

ART UNIT 1642, ROOM 3A24, MB 3C18

272-0830

cancelled

\*\*\*\*\*

## STAFF USE ONLY

Searcher: Arnold  
Searcher Phone: 2-2532  
Date Searcher Picked up: 2/25/05  
Date Completed: 3/3/05  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

## Type of Search

NA Sequence: # \_\_\_\_\_  
AA Sequence: # 1  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

## Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 07:50:03 ; Search time 70 Seconds  
(without alignments)  
872.973 Million cell updates/sec

Title: US-10-099-791E-2  
Perfect score: 878  
Sequence: 1 MASRSRRLLLLLSCLAKTGV.....LTWSSNECKRQHFLCKYRP 158

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A\_Geneseq\_16Dec04:.\*  
1: Geneseqp19808:.\*  
2: Geneseqp19908:.\*  
3: Geneseqp20008:.\*  
4: Geneseqp20018:.\*  
5: Geneseqp20028:.\*  
6: Geneseqp20038:.\*  
7: Geneseqp20038b:.\*  
8: Geneseqp20048:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878	100.0	158	2 AAW12691	AAW12691 Human col
2	878	100.0	158	2 AAW37929	AAW37929 A human R
3	878	100.0	158	2 AAW37866	AAW37866 Human pro
4	878	100.0	158	2 AAW84274	AAW84274 Protein e
5	878	100.0	158	3 AAB12900	AAB12900 Human col
6	878	100.0	158	4 AAB74934	AAB74934 Human TSA
7	878	100.0	158	4 AAM24519	AAM24519 C880P sim
8	878	100.0	158	4 AAM24520	AAM24520 C880P sim
9	878	100.0	158	4 AAM24521	AAM24521 C880P sim
10	878	100.0	158	4 AAM24517	AAM24517 CSI-152 c
11	878	100.0	158	4 AAM24518	AAM24518 C880P sim
12	878	100.0	158	5 AAE29829	AAE29829 Human REG
13	878	100.0	158	5 AAB78993	AAB78993 Human REG
14	878	100.0	158	6 ABR58551	ABR58551 Human can
15	878	100.0	158	6 ABP56022	ABP56022 Human REG
16	878	100.0	158	6 ADA83832	ADA83832 Human REG
17	878	100.0	158	6 ABP55366	ABP55366 Human col
18	878	100.0	158	6 ABP55370	ABP55370 Human col
19	878	100.0	158	6 ABP55369	ABP55369 Human col
20	878	100.0	158	6 ABP55367	ABP55367 Human col
21	878	100.0	158	6 ABP55368	ABP55368 Human col
22	878	100.0	158	7 ADB80529	ADB80529 Ovarian c
23	878	100.0	158	7 ADB75523	ADB75523 Prostate
24	878	100.0	158	7 ADF16659	ADF16659 Human alb
25	878	100.0	158	7 ADF56454	ADF56454 Human col

26	878	100.0	158	7 ADF85474	Adf85474 Human reg
27	878	100.0	158	7 ADH21872	Adh21872 Human Reg
28	878	100.0	158	7 ADN39463	Adn39463 Cancer/an
29	878	100.0	158	7 ADN39882	Adn39882 Cancer/an
30	878	100.0	158	7 ADN38820	Adn38820 Cancer/an
31	878	100.0	158	7 ADN39543	Adn39543 Cancer/an
32	878	100.0	158	8 ADI58211	Adi58211 Human reg
33	878	100.0	158	8 ADN35904	Adn35904 Human Reg
34	878	100.0	158	8 ADQ29679	Adq29679 Human col
35	878	100.0	158	8 ADQ80399	Adq80399 Regenerat
36	878	100.0	158	8 ADP55956	Adp55956 Human PRO
37	878	100.0	166	4 AAG75620	Aag75620 Human col
38	878	100.0	767	7 ADF16496	Adf16496 Human alb
39	878	100.0	767	7 ADH21795	Adh21795 Human alb
40	867	98.7	158	5 ABB99310	Abb99310 Amino aci
41	856	97.5	153	3 AAY92267	Aay92267 Human can
42	760	86.6	367	8 ADF73150	Adf73150 RELP-Pc f
43	695	79.2	122	6 ABP76307	Abp76307 Human GEN
44	307.5	35.0	134	5 ABP59097	Abp59097 Human zin
45	254.5	29.0	165	1 AAP81513	Aap81513 Sequence

ALIGNMENTS

RESULT 1  
AAW12691  
ID AAW12691 standard; protein; 158 AA.

XX AC AAW12691;  
XX DT 31-MAY-1997 (first entry)  
XX DE Human colon specific protein.  
XX KW Colon specific protein; colon cancer; metastasis; diagnosis; therapy;  
XX KW antibody; vaccine; agonist; antagonist.  
XX OS Homo sapiens.  
XX PN WO9639541-A1.  
XX PD 12-DEC-1996.  
XX PF 06-JUN-1995; 95WO-US007169.  
XX PR 06-JUN-1995; 95WO-US007169.  
XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Soppet DR, Li Y, Dillon PJ;  
XX DR WPI; 1997-043162/04.  
XX DR N-PSDB; AAT51784.  
XX PT New isolated colon specific gene - used to develop prods. for use in the  
XX diagnosis and treatment of colon disorders, partic. colon cancer.  
XX PS Claim 1; Page 53; 64pp; English.

CC A human colon specific protein (AAW12691) is a potential diagnostic  
CC marker for colon cancer. It is believed that the presence of active  
CC transcription of the colon specific gene in non-colon cells of a host is  
CC indicative of colon cancer metastases. The amino acid sequence of the  
CC colon specific protein was deduced from a cDNA clone (AAT51784) isolated  
CC from a human colon cancer cDNA library. Recombinant colon specific  
CC protein can be produced in transformed host (e.g. bacterial, insect)  
CC cells and used to develop prods. for the diagnosis and treatment of colon  
CC disorders, partic. colon cancer metastasis. Antibodies raised against the  
CC protein can be used to target cancer cells and as part of a colon cancer  
CC vaccine  
XX SQ Sequence 158 AA;

Query Match  
 Best Local Similarity 100.0%; Score 878; DB 2; Length 158;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 DB 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 QY 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120  
 DB 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120  
 QY 121 KSMGNGKCAEMSSNNFLTWSSNECNKRQHFCKYRP 158  
 DB 121 KSMGNGKCAEMSSNNFLTWSSNECNKRQHFCKYRP 158

RESULT 2  
 AAW37929  
 ID AAW37929 standard; protein; 158 AA.  
 AC AAW37929;  
 DT 21-AUG-1998 (first entry)  
 DE A human Reg I-gamma protein.  
 KW Human; Reg I-gamma protein; C-type lectin; Incyte clone 1310334;  
 KW regulation; cell growth; development; tumorigenesis; neurodegeneration;  
 KW inhibition; treatment; prevention; neoplasia; metastasis;  
 KW neurodegenerative change; Alzheimer's disease; Down's syndrome;  
 KW regeneration; pancreatic beta-cells; diabetes.  
 OS Homo sapiens.  
 PN WO9816640-A1.  
 PD 23-APR-1998.  
 PF 07-OCT-1997; 97WO-US018174.  
 PR 11-OCT-1996; 96US-00729103.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Bandman O, Goli SK;  
 DR WPI; 1998-251287/22.  
 PT N-PSDB; AAV29156.  
 PS New isolated human Reg I-gamma protein - useful for developing products  
 for treating, e.g. diabetes, tumours or neuro-degenerative disease such  
 as Alzheimer's.

Claim 1; Fig 1; 72pp; English.  
 The present sequence represents a human Reg I-gamma protein, which  
 comprises a C-type lectin. The sequence was identified in Incyte clone  
 1310334. Reg I-gamma protein is involved in regulation of cell growth and  
 development. Since the overexpression of reg proteins is associated with  
 tumorigenesis and neurodegeneration, inhibition of human Reg I-gamma  
 expression can be used for treating or preventing neoplasia or metastasis  
 and neurodegenerative changes associated with Alzheimer's disease and  
 other disorders of the central nervous system, e.g. Down's syndrome. Reg  
 I-gamma can also be used in therapeutics to induce regeneration of  
 pancreatic beta-cells in the treatment of diabetes. The products can also  
 be used for detection for, e.g. expression of REG I-gamma, diagnosis and  
 drug screen

Query Match  
 Best Local Similarity 100.0%; Score 878; DB 2; Length 158;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 DB 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 QY 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120  
 DB 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120

Query Match  
 Best Local Similarity 100.0%; Score 878; DB 2; Length 158;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 DB 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 QY 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120  
 DB 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120

Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 DB 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 QY 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120  
 DB 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120  
 QY 121 KSMGNGKCAEMSSNNFLTWSSNECNKRQHFCKYRP 158  
 DB 121 KSMGNGKCAEMSSNNFLTWSSNECNKRQHFCKYRP 158

RESULT 3  
 AAW37866  
 ID AAW37866 standard; protein; 158 AA.  
 AC AAW37866;  
 DT 10-AUG-1998 (first entry)  
 DE Human protein comprising secretory signal amino acid sequence 3.  
 KW Human protein; secretory signal; nutritional source; cytokine; immunity;  
 KW haematopoiesis; activin; inhibin; tumour; chemotactic; chemokinetic;  
 KW thrombolytic; anti-inflammatory; inhibition; stomach cancer cell.  
 OS Homo sapiens.  
 PN WO9811217-A2.  
 PD 19-MAR-1998.  
 PF 12-SEP-1997; 97WO-JP003239.  
 PR 13-SEP-1996; 96JP-00243060.  
 PA (SAGA) SAGAMI CHEM RES CENTRE.  
 PI (PROT-) PROTEGENE INC.  
 PI Kato S, Sekine S, Yamaguchi T, Kobayashi M;  
 DR WPI; 1998-207380/18.  
 PT N-PSDB; AAV29035, AAV29036.  
 PS Human proteins with secretory signal sequences - used to treat immune  
 deficiencies, infections, tumours, and haematopoietic disorders, etc.

Claim 1; Page 67-68; 131pp; English.  
 This is the amino acid sequence of a novel human protein comprising a  
 secretory signal isolated from stomach cancer cells. Its proteins can be  
 used as nutritional sources or supplements. The proteins may also have  
 cytokine functions, immune modulating functions, haematopoiesis  
 regulating activity, activin/inhibin regulating activity,  
 chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 receptor/ligand activity, anti-inflammatory activity, tumour inhibition  
 activity

Query Match  
 Best Local Similarity 100.0%; Score 878; DB 2; Length 158;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 DB 1 MASRSMRLLLLSCLAKTGVLDGIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
 QY 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120  
 DB 61 YNGAHLASTLSLKEASTIAEYISGYQSORPIMTGLHDPKQKQOWIDGAMYLRSWSG 120

Db 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWGLHDPKQKQOWIDGAMVLYRSWSG 120  
 QY 121 KSMGKNGKCAEMSSNNFLTWSSNECKRQHFCKYRP 158  
 Db 121 KSMGKNGKCAEMSSNNFLTWSSNECKRQHFCKYRP 158

RESULT 4  
 AAW84274  
 ID AAW84274 standard; protein; 158 AA.  
 XX  
 AC AAW84274;  
 XX  
 DT 25-MAR-1999 (first entry)  
 XX  
 DE Protein encoded by a human colon specific gene.  
 XX  
 KW Human; colon specific gene; diagnosis; colon disorder; colon cancer;  
 KW viability; colon cancer cell.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US861494-A.  
 XX  
 PD 19-JAN-1999.  
 XX  
 PF 06-JUN-1995; 95US-00468413.  
 XX  
 PR 06-JUN-1995; 95US-00468413.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Dillion PJ, Soppet DR, Li Y;  
 XX  
 PS WPI; 1999-130432/11.  
 DR N-PSDB; AAX03195.

Isolated human colon specific gene - used to develop products for the diagnosis and treatment of disorders of the colon, e.g. colon cancer and metastases.  
 Claim 1; Fig 1A-C; 20pp; English.  
 The present sequence is encoded by a human colon specific gene. The nucleic acid sequence can be used to develop products for the diagnosis of a disorder of the colon, e.g. colon cancer or metastases. The products can also be used to screen for agonists or antagonists for the colon polypeptides. The antagonists may be used to treat colon cancer, since they interact with the function of colon specific polypeptides to inhibit functions which are necessary for the viability of colon cancer cells. The products can also be used for the production of antibodies and for the identification of receptors for the polypeptides

XX  
 SQ Sequence 158 AA;  
 Query Match 100.0%; Score 878; DB 2; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGYFKLRNWSDALECSQ 60  
 Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGYFKLRNWSDALECSQ 60  
 QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWGLHDPKQKQOWIDGAMVLYRSWSG 120  
 Db 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWGLHDPKQKQOWIDGAMVLYRSWSG 120  
 QY 121 KSMGKNGKCAEMSSNNFLTWSSNECKRQHFCKYRP 158  
 Db 121 KSMGKNGKCAEMSSNNFLTWSSNECKRQHFCKYRP 158

RESULT 5  
 AAB12900  
 ID AAB12900 standard; protein; 158 AA.  
 XX  
 AC AAB12900;  
 XX  
 DT 14-NOV-2000 (first entry)  
 XX  
 DE Human colon specific protein sequence.  
 XX  
 KW Human; colon specific; colon cancer; metastasis; diagnose; treatment;  
 KW cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6080722-A.  
 XX  
 PD 27-JUN-2000.  
 XX  
 PF 29-SEP-1998; 98US-00162508.  
 XX  
 PR 06-JUN-1995; 95US-00468413.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Dillion PJ, Soppet DR, Li Y;  
 XX  
 PS WPI; 2000-464035/40.  
 DR N-PSDB; AAA62951.

Novel human colon specific polypeptides and polynucleotides for diagnosis and treatment of colon cancer, for screening compounds which interact with polypeptide, for synthesis of DNA and manufacture of DNA vectors.  
 Claim 1; Fig 1; 20pp; English.

XX  
 SQ Sequence 158 AA;  
 Query Match 100.0%; Score 878; DB 3; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGYFKLRNWSDALECSQ 60  
 Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGYFKLRNWSDALECSQ 60  
 QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWGLHDPKQKQOWIDGAMVLYRSWSG 120  
 Db 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWGLHDPKQKQOWIDGAMVLYRSWSG 120  
 QY 121 KSMGKNGKCAEMSSNNFLTWSSNECKRQHFCKYRP 158  
 Db 121 KSMGKNGKCAEMSSNNFLTWSSNECKRQHFCKYRP 158

RESULT 6  
 AAB74934  
 ID AAB74934 standard; protein; 158 AA.  
 XX  
 AC AAB74934;  
 XX  
 DT 27-JUN-2001 (first entry)  
 XX  
 PS Human TSA7005 protein SEQ ID NO:1.

KW Human; TSA7005; Reg; pancreatic beta cell growth; hypoglycaemic;  
 KW diagnosis.  
 XX Homo sapiens.  
 OS JP2001025389-A.  
 PN 30-JAN-2001.  
 PD 15-JUL-1999; 99JP-00201279.  
 PF 15-JUL-1999; 99JP-00201279.  
 XX (SAKA ) OTSUKA PHARM CO LTD.  
 PA WPI; 2001-303742/32.  
 DR N-PSDB; AAF82117, AAF82118.  
 XX TSA7005 gene, encoding a polypeptide useful for the diagnosis and  
 PT treatment of diseases associated with its expression.  
 PT Claim 1; Page 23; 25pp; Japanese.  
 XX The present sequence represents a human TSA7005 protein which shares 32%  
 CC homology with human and mouse Reg proteins, and 34% homology with the rat  
 CC Reg protein. TSA7005 has pancreatic beta cell growth activity and  
 CC hypoglycaemic activity. The TSA7005 protein can be used for the diagnosis  
 CC and treatment of diseases associated with the gene and its expression  
 CC Product  
 XX Sequence 158 AA;  
 SQ Query Match 100.0%; Score 878; DB 4; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASRSNRLILLLSCLAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNWSDALEQCS 60  
 DB 1 MASRSNRLILLLSCLAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNWSDALEQCS 60  
 QY 61 YGNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
 DB 61 YGNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
 QY 121 KSMGNGKHCAEMSSNNFLTWSSNECNKROHFLCKYRP 158  
 DB 121 KSMGNGKHCAEMSSNNFLTWSSNECNKROHFLCKYRP 158  
 RESULT 7  
 ID AAM24519 standard; protein; 158 AA.  
 XX AAM24519;  
 AC AAM24519;  
 XX 12-OCT-2001 (first entry)  
 XX C880P similar amino acid sequence (GENESEQ W37866).  
 DE Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;  
 KW gene therapy; vaccine; colonic cancer.  
 XX Homo sapiens.  
 OS WO200149716-A2.  
 PN 12-JUL-2001.  
 PD 29-DEC-2000; 2000WO-US035596.  
 PF 30-DEC-1999; 99US-00476296.  
 PR 10-JAN-2000; 2000US-00480321.  
 PR 15-FEB-2000; 2000US-00504629.

PR 06-MAR-2000; 2000US-00519444.  
 PR 19-MAY-2000; 2000US-00575251.  
 PR 29-JUN-2000; 2000US-00609448.  
 PR 28-AUG-2000; 2000US-00649811.  
 XX (CORI-) CORIXA CORP.  
 XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;  
 PI King GE, Wang T, Jiang Y;  
 XX WPI; 2001-441847/47.  
 DR Colon tumor associated proteins and nucleic acids useful for the  
 XX prevention, diagnosis and treatment of colonic cancer.  
 PT Claim 2; Page 467-468; 472pp; English.  
 XX The present invention describes colon tumour associated proteins (I) and  
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.  
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and  
 CC (II) may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate colon tumour associated protein (TCAP)  
 CC expression, such as colonic cancer. For example, (I) and (II) may be used  
 CC to treat disorders associated with decreased expression by rectifying  
 CC mutations or deletions in a patient's genome that affect the activity of  
 CC TCAPs by expressing inactive proteins or to supplement the patients own  
 CC production of them. Additionally, (II) may be used to produce the TCAP  
 CC proteins, by inserting the nucleic acids into a host cell culturing the  
 CC cell to express the protein. (II) and its complementary sequences may  
 CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)  
 CC and hybridisation assays to detect and quantitate the presence of similar  
 CC nucleic acids in samples, and therefore which patients may be in need of  
 CC restorative therapy. (II) may also be used as antigens in the production  
 CC of antibodies against TCAPs and in assays to identify modulators of TCAP  
 CC expression and activity. Anti-(I) antibodies and antagonists may also be  
 CC used to down regulate TCAP expression and activity. The anti-(I)  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbent assay  
 CC (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent  
 CC nucleotide and amino acid sequences given in the exemplification of the  
 CC present invention  
 XX Sequence 158 AA;  
 SQ Query Match 100.0%; Score 878; DB 4; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASRSNRLILLLSCLAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNWSDALEQCS 60  
 DB 1 MASRSNRLILLLSCLAKTGVLDIIMRPSCAPGFYHKSNCYGYFRKLRNWSDALEQCS 60  
 QY 61 YGNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
 DB 61 YGNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
 QY 121 KSMGNGKHCAEMSSNNFLTWSSNECNKROHFLCKYRP 158  
 DB 121 KSMGNGKHCAEMSSNNFLTWSSNECNKROHFLCKYRP 158  
 RESULT 8  
 ID AAM24520 standard; protein; 158 AA.  
 XX AAM24520;  
 AC AAM24520;  
 XX 12-OCT-2001 (first entry)  
 XX C880P similar amino acid sequence (GENESEQ W37929).  
 DE Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;  
 KW gene therapy; vaccine; colonic cancer.



XX OS Homo sapiens.  
 XX OS AAM24521  
 XX ID AAM24521 standard; protein; 158 AA.  
 XX AC AAM24521;  
 XX AC AAM24521;  
 XX PD 12-JUL-2001.  
 XX PF 29-DEC-2000; 2000WO-US035596.  
 XX PR 30-DEC-1999; 99US-00476296.  
 XX PR 10-JAN-2000; 2000US-00480321.  
 XX PR 15-FEB-2000; 2000US-00504629.  
 XX PR 06-MAR-2000; 2000US-00519444.  
 XX PR 19-MAY-2000; 2000US-00575251.  
 XX PR 29-JUN-2000; 2000US-00609448.  
 XX PR 28-AUG-2000; 2000US-00649811.  
 XX PA (CORI-) CORIXA CORP.  
 XX PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;  
 XX PI King GE, Wang T, Jiang Y;  
 XX DR WPI; 2001-441847/47.  
 XX PT Colon tumor associated proteins and nucleic acids useful for the  
 XX PT prevention, diagnosis and treatment of colonic cancer.  
 XX PS Claim 2; Page 468; 472pp; English.  
 XX CC The present invention describes colon tumour associated proteins (I) and  
 XX CC the polynucleotides (II) that encode them. (I) have cytostatic activity.  
 XX CC (I) and (II) can be used in gene therapy and vaccine production. (I) and  
 XX CC (II) may be used in the prevention, diagnosis and treatment of diseases  
 XX CC associated with inappropriate colon tumour associated protein (TCAP)  
 XX CC expression, such as colonic cancer. For example, (I) and (II) may be used  
 XX CC to treat disorders associated with decreased expression by rectifying  
 XX CC mutations or deletions in a patient's genome that affect the activity of  
 XX CC TCAPs by expressing inactive proteins or to supplement the patients own  
 XX CC production of them. Additionally, (II) may be used to produce the TCAP  
 XX CC proteins, by inserting the nucleic acids into a host cell culturing the  
 XX CC cell to express the protein. (II) and its complementary sequences may  
 XX CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)  
 XX CC and hybridisation assays to detect and quantitate the presence of similar  
 XX CC nucleic acids in samples, and therefore which patients may be in need of  
 XX CC restorative therapy. (I) may also be used as antigens in the production  
 XX CC of antibodies against TCAPs and in assays to identify modulators of TCAP  
 XX CC expression and activity. Anti-(I) antibodies and antagonists may also be  
 XX CC used to down regulate TCAP expression and activity. The anti-(I)  
 XX CC antibodies may also be used as diagnostic agents for detecting the  
 XX CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbent assay  
 XX CC (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent  
 XX CC nucleotide and amino acid sequences given in the exemplification of the  
 XX CC present invention  
 XX SQ Sequence 158 AA;  
 Query Match 100.0%; Score 878; DB 4; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALEECQS 60  
 Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALEECQS 60  
 Qy 61 YGNGAHLASILSLKEASTIAEYISGYQSRQSPITWGLHDHPKQKQOWMDGAWLYRSWSG 120  
 Db 61 YGNGAHLASILSLKEASTIAEYISGYQSRQSPITWGLHDHPKQKQOWMDGAWLYRSWSG 120  
 Qy 121 KSMGKXKHCAMSSNNNPLTWSSNECKNRQHFLCKYRP 158  
 Db 121 KSMGKXKHCAMSSNNNPLTWSSNECKNRQHFLCKYRP 158

RESULT 9  
 AAM24521  
 ID AAM24521 standard; protein; 158 AA.  
 XX AC AAM24521;  
 XX AC AAM24521;  
 XX DT 12-OCT-2001 (first entry)  
 XX DE C880P similar amino acid sequence (GENESEQ W84274).  
 XX KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;  
 XX KW gene therapy; vaccine; colonic cancer.  
 XX OS Homo sapiens.  
 XX PN WO200149716-A2.  
 XX PD 12-JUL-2001.  
 XX PF 29-DEC-2000; 2000WO-US035596.  
 XX PR 30-DEC-1999; 99US-00476296.  
 XX PR 10-JAN-2000; 2000US-00480321.  
 XX PR 15-FEB-2000; 2000US-00504629.  
 XX PR 06-MAR-2000; 2000US-00519444.  
 XX PR 19-MAY-2000; 2000US-00575251.  
 XX PR 29-JUN-2000; 2000US-00609448.  
 XX PR 28-AUG-2000; 2000US-00649811.  
 XX PA (CORI-) CORIXA CORP.  
 XX PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;  
 XX PI King GE, Wang T, Jiang Y;  
 XX DR WPI; 2001-441847/47.  
 XX PT Colon tumor associated proteins and nucleic acids useful for the  
 XX PT prevention, diagnosis and treatment of colonic cancer.  
 XX PS Claim 2; Page 469; 472pp; English.  
 XX CC The present invention describes colon tumour associated proteins (I) and  
 XX CC the polynucleotides (II) that encode them. (I) have cytostatic activity.  
 XX CC (I) and (II) can be used in gene therapy and vaccine production. (I) and  
 XX CC (II) may be used in the prevention, diagnosis and treatment of diseases  
 XX CC associated with inappropriate colon tumour associated protein (TCAP)  
 XX CC expression, such as colonic cancer. For example, (I) and (II) may be used  
 XX CC to treat disorders associated with decreased expression by rectifying  
 XX CC mutations or deletions in a patient's genome that affect the activity of  
 XX CC TCAPs by expressing inactive proteins or to supplement the patients own  
 XX CC production of them. Additionally, (II) may be used to produce the TCAP  
 XX CC proteins, by inserting the nucleic acids into a host cell culturing the  
 XX CC cell to express the protein. (II) and its complementary sequences may  
 XX CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)  
 XX CC and hybridisation assays to detect and quantitate the presence of similar  
 XX CC nucleic acids in samples, and therefore which patients may be in need of  
 XX CC restorative therapy. (I) may also be used as antigens in the production  
 XX CC of antibodies against TCAPs and in assays to identify modulators of TCAP  
 XX CC expression and activity. Anti-(I) antibodies and antagonists may also be  
 XX CC used to down regulate TCAP expression and activity. The anti-(I)  
 XX CC antibodies may also be used as diagnostic agents for detecting the  
 XX CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbent assay  
 XX CC (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent  
 XX CC nucleotide and amino acid sequences given in the exemplification of the  
 XX CC present invention  
 XX SQ Sequence 158 AA;  
 Query Match 100.0%; Score 878; DB 4; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALEECQS 60  
 Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALEECQS 60  
 Qy 61 YGNGAHLASILSLKEASTIAEYISGYQSRQSPITWGLHDHPKQKQOWMDGAWLYRSWSG 120  
 Db 61 YGNGAHLASILSLKEASTIAEYISGYQSRQSPITWGLHDHPKQKQOWMDGAWLYRSWSG 120  
 Qy 121 KSMGKXKHCAMSSNNNPLTWSSNECKNRQHFLCKYRP 158  
 Db 121 KSMGKXKHCAMSSNNNPLTWSSNECKNRQHFLCKYRP 158

Db 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRKLRNWSDALEQCS 60  
QY 61 YNGAHLASILSLKEASTIAEYISGYQSRQPIWIGLHDPKQKQOWIDGAMVLYRSMSG 120  
Db 61 YNGAHLASILSLKEASTIAEYISGYQSRQPIWIGLHDPKQKQOWIDGAMVLYRSMSG 120  
QY 121 KSMGNGKHCAMSSNNFLTWSSNECKRQHFLCKYRP 158  
Db 121 KSMGNGKHCAMSSNNFLTWSSNECKRQHFLCKYRP 158

RESULT 10  
AAM24517  
ID AAM24517 standard; protein; 158 AA.  
XX  
AC AAM24517;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE CS1-152 clone predicted amino acid sequence.  
XX  
DE Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;  
KW Gene therapy; vaccine; colonic cancer.  
KW  
OS Homo sapiens.  
EN WO200149716-A2.  
XX  
PD 12-JUL-2001.  
XX  
PF 29-DEC-2000; 2000WO-US035596.  
XX  
PR 30-DEC-1999; 99US-00476296.  
PR 10-JAN-2000; 2000US-00480321.  
PR 15-FEB-2000; 2000US-00504629.  
PR 06-MAR-2000; 2000US-00519444.  
PR 19-MAY-2000; 2000US-00575251.  
PR 29-JUN-2000; 2000US-00609448.  
PR 28-AUG-2000; 2000US-00649811.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;  
PI King GE, Wang T, Jiang Y;  
XX  
DR WPI; 2001-441847/47.  
XX  
PT Colon tumor associated proteins and nucleic acids useful for the  
PT prevention, diagnosis and treatment of colonic cancer.  
XX  
PS Claim 2; Page 463; 472pp; English.  
XX  
CC The present invention describes colon tumour associated proteins (I) and  
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.  
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and  
CC (II) may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate colon tumour associated protein (TCAP)  
CC expression, such as colonic cancer. For example, (I) and (II) may be used  
CC to treat disorders associated with decreased expression by rectifying  
CC mutations or deletions in a patient's genome that affect the activity of  
CC TCAPs by expressing inactive proteins or to supplement the patients own  
CC production of them. Additionally, (II) may be used to produce the TCAP  
CC proteins, by inserting the nucleic acids into a host cell culturing the  
CC cell to express the protein. (II) and its complementary sequences may  
CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)  
CC and hybridisation assays to detect and quantitate the presence of similar  
CC nucleic acids in samples, and therefore which patients may be in need of  
CC resorcinative therapy. (I) may also be used as antigens in the production  
CC of antibodies against TCAPs and in assays to identify modulators of TCAP  
CC expression and activity. Anti-(I) antibodies and antagonists may also be  
CC used to down regulate TCAP expression and activity. The anti-(I)  
CC antibodies may also be used as diagnostic agents for detecting the

CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay  
CC (ELISA)). AA128460 to AA129512 and AAM24494 to AAM24523 represent  
CC nucleotide and amino acid sequences given in the exemplification of the  
CC present invention  
XX  
SQ Sequence 158 AA;  
Query Match 100.0%; Score 878; DB 4; Length 158;  
Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRKLRNWSDALEQCS 60  
Db 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFRKLRNWSDALEQCS 60  
QY 61 YNGAHLASILSLKEASTIAEYISGYQSRQPIWIGLHDPKQKQOWIDGAMVLYRSMSG 120  
Db 61 YNGAHLASILSLKEASTIAEYISGYQSRQPIWIGLHDPKQKQOWIDGAMVLYRSMSG 120  
QY 121 KSMGNGKHCAMSSNNFLTWSSNECKRQHFLCKYRP 158  
Db 121 KSMGNGKHCAMSSNNFLTWSSNECKRQHFLCKYRP 158

RESULT 11  
AAM24518  
ID AAM24518 standard; protein; 158 AA.  
XX  
AC AAM24518;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE C880P similar amino acid sequence (GENESEQ W12691).  
XX  
DE Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;  
KW Gene therapy; vaccine; colonic cancer.  
KW  
OS Homo sapiens.  
EN WO200149716-A2.  
XX  
PD 12-JUL-2001.  
XX  
PF 29-DEC-2000; 2000WO-US035596.  
XX  
PR 30-DEC-1999; 99US-00476296.  
PR 10-JAN-2000; 2000US-00480321.  
PR 15-FEB-2000; 2000US-00504629.  
PR 06-MAR-2000; 2000US-00519444.  
PR 19-MAY-2000; 2000US-00575251.  
PR 29-JUN-2000; 2000US-00609448.  
PR 28-AUG-2000; 2000US-00649811.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;  
PI King GE, Wang T, Jiang Y;  
XX  
DR WPI; 2001-441847/47.  
XX  
PT Colon tumor associated proteins and nucleic acids useful for the  
PT prevention, diagnosis and treatment of colonic cancer.  
XX  
PS Claim 2; Page 467; 472pp; English.  
XX  
CC The present invention describes colon tumour associated proteins (I) and  
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.  
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and  
CC (II) may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate colon tumour associated protein (TCAP)  
CC expression, such as colonic cancer. For example, (I) and (II) may be used  
CC to treat disorders associated with decreased expression by rectifying  
CC mutations or deletions in a patient's genome that affect the activity of  
CC TCAPs by expressing inactive proteins or to supplement the patients own  
CC production of them. Additionally, (II) may be used to produce the TCAP  
CC proteins, by inserting the nucleic acids into a host cell culturing the  
CC cell to express the protein. (II) and its complementary sequences may  
CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)  
CC and hybridisation assays to detect and quantitate the presence of similar  
CC nucleic acids in samples, and therefore which patients may be in need of  
CC resorcinative therapy. (I) may also be used as antigens in the production  
CC of antibodies against TCAPs and in assays to identify modulators of TCAP  
CC expression and activity. Anti-(I) antibodies and antagonists may also be  
CC used to down regulate TCAP expression and activity. The anti-(I)  
CC antibodies may also be used as diagnostic agents for detecting the

CC TCAPs by expressing inactive proteins or to supplement the patients own  
 CC production of them. Additionally, (II) may be used to produce the TCAP  
 CC proteins, by inserting the nucleic acids into a host cell culturing the  
 CC cell to express the protein. (II) and its complementary sequences may  
 CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)  
 CC and hybridization assays to detect and quantitate the presence of similar  
 CC nucleic acids in samples, and therefore which patients may be in need of  
 CC restorative therapy. (I) may also be used as antigens in the production  
 CC of antibodies against TCAPs and in assays to identify modulators of TCAP  
 CC expression and activity. Anti-(I) antibodies and antagonists may also be  
 CC used to down regulate TCAP expression and activity. The anti-(I)  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay  
 CC (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent  
 CC nucleotide and amino acid sequences given in the exemplification of the  
 CC present invention

XX Sequence 158 AA;  
 SQ

Query Match 100.0%; Score 878; DB 4; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFKLRNWSDALEECQS 60  
 DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFKLRNWSDALEECQS 60  
 QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWGLHDPKROQWOWIDGAMLYRSWSG 120  
 DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWGLHDPKROQWOWIDGAMLYRSWSG 120  
 QY 121 KSMGKNKHCAEMSSNNFLTWSSNECNKROHFLCKYRP 158  
 DB 121 KSMGKNKHCAEMSSNNFLTWSSNECNKROHFLCKYRP 158

## RESULT 12

AAE29829  
 ID AAE29829 standard; protein; 158 AA.

XX AAE29829;

XX 24-FEB-2003 (first entry)

XX Human REG-like protein (RELIP).

XX Human; REG-like protein; RELIP; tumour; cancer; therapy; chromosome 1.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT Protein /label= Signal\_peptide

FT 23..158

FT /note= "Human mature REG-like protein"

XX EP1241269-A2.

XX 18-SEP-2002.

XX 15-MAR-2002; 2002EP-00251876.

XX 16-MAR-2001; 2001US-0276414P.

XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.

XX Heiskala M;

XX WPI; 2002-684095/74.

XX N-PSDB; AAD47239.

XX Detecting the presence of a tumor comprises detecting the concentration  
 FT of a Reg Like Protein or the presence or quantity of a nucleic acid

PT encoding it.

XX Claim 1; Page 13-14; 26pp; English.

XX The invention relates to a method for detecting REG-like protein (RELIP)  
 CC and its nucleic acid sequence. The method is useful for detecting the  
 CC presence of a tumour. Kits comprising an antibody specific for RELIP and  
 CC reagents for detecting the antibody, or a nucleic acid complementary to a  
 CC portion of a nucleic acid encoding RELIP, are useful for identifying the  
 CC presence of cancer, characterise the cancer, or monitor the course of  
 CC treatment of cancer. The present sequence is human RELP protein used to  
 CC illustrate the method of the invention. Human RELP gene is located at  
 CC chromosome 1

XX Sequence 158 AA;

Query Match 100.0%; Score 878; DB 5; Length 158;

Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFKLRNWSDALEECQS 60

DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFKLRNWSDALEECQS 60

QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWGLHDPKROQWOWIDGAMLYRSWSG 120

DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWGLHDPKROQWOWIDGAMLYRSWSG 120

QY 121 KSMGKNKHCAEMSSNNFLTWSSNECNKROHFLCKYRP 158

DB 121 KSMGKNKHCAEMSSNNFLTWSSNECNKROHFLCKYRP 158

## RESULT 13

ABB78993

ID ABB78993 standard; protein; 158 AA.

XX ABB78993;

XX 02-AUG-2002 (first entry)

XX Human Reg IV protein sequence SEQ ID NO:4471.

XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
 genetic analysis; diagnostic; antisense therapy.

XX Homo sapiens.

XX WO200229086-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US030732.

XX 02-OCT-2000; 2000US-0237271P.

XX (FARB ) BAYER CORP.

XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;

XX Thiagalingam A, Lewis ME;

XX WPI; 2002-426115/45.

XX N-PSDB; ABQ60776.

XX New isolated nucleic acid that is differentially expressed in cancer  
 PT tissues useful for determining the presence of colon cancer in a cell or  
 PT tissue type, and in antisense therapy.

XX Claim 5; Fig 3; 795pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
 CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins  
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be

CC used in antineoplastic therapy. An antibody immunoreactive with a polypeptide  
CC encoded by (I) is useful for detecting cancer in a patient sample, and  
CC for detecting the presence or absence of a polynucleotide encoded by a  
CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived  
CC from (II) can be used for determining the presence of a nucleic acid which  
CC hybridizes to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence of  
CC colon cancer in a cell or tissue type, for determining the presence or  
CC state of other type of cancer, in antisense therapy, to generate  
CC macroarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (I) can be used to raise  
CC antibodies, and to screen for peptide analogues and antagonists  
XX  
XX Sequence 158 AA;

Query Match 100.0%; Score 878; DB 5; Length 158;  
Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
QY 61 YNGAHLASLILSKEASTIAEYISGYQRSQPIWIGLHDPKQKQWOWIDGAMLYRSWSG 120  
DB 61 YNGAHLASLILSKEASTIAEYISGYQRSQPIWIGLHDPKQKQWOWIDGAMLYRSWSG 120  
QY 121 KSMGKNKHCAPMSNNNFLTWSNECNKROHFLCKYRP 158  
DB 121 KSMGKNKHCAPMSNNNFLTWSNECNKROHFLCKYRP 158

RESULT 14  
ABRS551  
ID ABR58551 standard; protein; 158 AA.  
XX  
AC ABR58551;  
XX  
DT 09-JUL-2003 (first entry)  
XX  
DE Human cancer related protein SEQ ID NO:208.

XX Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;  
KW heart disease; atherosclerosis; endometriosis.  
XX  
XX Homo sapiens.

XX WO2003025138-A2.  
XX  
XX 27-MAR-2003.  
XX  
XX 17-SEP-2002; 2002WO-US029560.  
XX  
XX 17-SEP-2001; 2001US-0323469P.  
XX  
XX 20-SEP-2001; 2001US-0323887P.  
XX  
XX 13-NOV-2001; 2001US-0350666P.  
XX  
XX 08-FEB-2002; 2002US-0355145P.  
XX  
XX 08-FEB-2002; 2002US-0355257P.  
XX  
XX 12-APR-2002; 2002US-0372246P.  
XX  
XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;  
PI Zlotnik A;  
XX  
XX WPI; 2003-354600/33.  
XX  
XX N-PSDB; ACC72672.

XX New genes that are up-regulated or down-regulated in cancers, useful as  
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as  
PT therapeutic targets for screening drugs for treating these diseases.  
XX  
XX

PS Claim 12; Page 737; 767pp; English.  
XX  
CC The present invention describes an isolated nucleic acid molecule, which  
CC comprises the sequence of any of the genes that are up-regulated or down-  
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in  
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer  
CC related gene nucleotide sequences which encode the proteins given in  
CC ABR58521 to ABR58709. Also described: (1) determining the presence or  
CC absence of a pathological cell in a patient; (2) an expression vector  
CC comprising a nucleic acid molecule described above; (3) a host cell  
CC of the nucleic acid; (4) an isolated polypeptide, which is encoded by  
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide  
CC of (4); (6) specifically targeting a compound to a pathological cell in a  
CC patient by administering to the patient the antibody above; and (7) a  
CC drug screening assay. The nucleic acid is useful as diagnostic markers or  
CC therapeutic targets. In particular, the nucleic acid is useful for  
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,  
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,  
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,  
CC atherosclerosis and endometriosis. The nucleic acid is also useful in  
CC drug screening, particularly for identifying agents for treating these  
CC pathologies  
XX  
XX Sequence 158 AA;

Query Match 100.0%; Score 878; DB 6; Length 158;  
Best Local Similarity 100.0%; Pred. No. 6.8e-83;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLNWSDAELECS 60  
QY 61 YNGAHLASLILSKEASTIAEYISGYQRSQPIWIGLHDPKQKQWOWIDGAMLYRSWSG 120  
DB 61 YNGAHLASLILSKEASTIAEYISGYQRSQPIWIGLHDPKQKQWOWIDGAMLYRSWSG 120  
QY 121 KSMGKNKHCAPMSNNNFLTWSNECNKROHFLCKYRP 158  
DB 121 KSMGKNKHCAPMSNNNFLTWSNECNKROHFLCKYRP 158

RESULT 15  
ABP56022  
ID ABP56022 standard; protein; 158 AA.  
XX  
AC ABP56022;  
XX  
DT 26-FEB-2003 (first entry)  
XX  
XX Human REG-like protein (RELp) SEQ ID NO:2.

XX Human; REG-like protein; RELp; immunoglobulin derived protein; Ig;  
KW immunoglobulin; cytostatic; Ig agonist; immunoglobulin agonist; cancer;  
XX protein therapy; RELp human Ig derived protein; chromosome ip12-13.1.  
XX  
XX Homo sapiens.

XX Key Location/Qualifiers  
FT Peptide 1..26 /label= signal  
FT Protein 27..158 /label= RELp  
FT  
XX WO200274916-A2.

XX 26-SEP-2002.  
XX  
XX 14-MAR-2002; 2002WO-US007945.  
XX  
XX 16-MAR-2001; 2001US-0276305P.  
XX  
XX (CENZ ) CENTOCOR INC.





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 07:50:08 ; Search time 22 Seconds  
(without alignments)  
536.116 Million cell updates/sec

Title: US-10-099-791E-2  
Perfect score: 878  
Sequence: 1 MASRSWRLLLLSCLAKTV.....LTWSSNECKRQHFLCKYRP 158

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	878	100.0	158	2	US-08-729-103-1
2	878	100.0	158	2	US-08-468-413-2
3	878	100.0	158	3	US-09-162-508-2
4	878	100.0	158	5	PCT-US95-07169-2
5	695	79.2	122	4	US-09-621-976-4812
6	254.5	29.0	165	2	US-08-401-530A-7
7	254.5	29.0	165	2	US-08-729-103-3
8	254.5	29.0	165	2	US-08-709-662-7
9	240	27.3	166	2	US-08-729-103-4
10	240	27.3	174	4	US-09-949-016-10686
11	231	26.3	174	2	US-08-401-530A-5
12	231	26.3	174	2	US-08-709-662-5
13	228.5	26.0	174	2	US-08-401-530A-6
14	228.5	26.0	174	2	US-08-709-662-6
15	224	25.5	144	4	US-09-949-016-10685
16	224	25.5	166	4	US-09-949-016-6286
17	223	25.4	117	6	5514582-15
18	223	25.4	117	6	5514582-15
19	221	25.2	175	2	US-08-464-637-2
20	221	25.2	175	2	US-08-401-530A-4
21	221	25.2	175	2	US-08-709-662-4
22	221	25.2	175	2	US-08-822-261-3
23	221	25.2	175	4	US-09-226-852-3
24	221	25.2	183	4	US-09-949-016-10537
25	218	24.8	174	1	US-07-778-156-7
26	218	24.8	174	2	US-08-822-261-4
27	218	24.8	174	2	US-08-422-166-7

28	218	24.8	174	4	US-09-226-852-4	Sequence 4, Appli
29	208	23.7	175	2	US-08-401-530A-3	Sequence 3, Appli
30	208	23.7	175	2	US-08-709-662-3	Sequence 3, Appli
31	204.5	23.3	130	1	US-07-893-929A-7	Sequence 7, Appli
32	204.5	23.3	130	5	PCT-US92-10344-7	Sequence 7, Appli
33	204	23.2	175	2	US-08-822-261-1	Sequence 1, Appli
34	204	23.2	175	4	US-09-226-852-1	Sequence 1, Appli
35	198.5	22.6	132	1	US-07-893-929A-5	Sequence 5, Appli
36	198.5	22.6	132	5	PCT-US92-10344-5	Sequence 5, Appli
37	198.5	22.6	134	1	US-07-893-929A-2	Sequence 2, Appli
38	198.5	22.6	134	5	PCT-US92-10344-2	Sequence 2, Appli
39	196	22.3	125	1	US-07-893-929A-3	Sequence 3, Appli
40	196	22.3	125	5	PCT-US92-10344-3	Sequence 3, Appli
41	194.5	22.2	131	1	US-07-893-929A-1	Sequence 1, Appli
42	194.5	22.2	131	5	PCT-US92-10344-1	Sequence 1, Appli
43	193	22.0	912	5	PCT-US95-03747-2	Sequence 2, Appli
44	190	21.6	174	2	US-08-401-530A-2	Sequence 2, Appli
45	190	21.6	174	2	US-08-709-662-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-729-103-1  
; Sequence 1, Application US/08729103  
; Patent No. 5837841  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/729,103  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0138 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 158 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; LIBRARY: COLNFET02  
; CLONE: 1310334  
US-08-729-103-1

Query Match 100.0%; Score 878; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 5.9e-88;  
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60  
 DB 1 MASRSMRLLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60  
 QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQWIDGAMLYYRSWSG 120  
 DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQWIDGAMLYYRSWSG 120  
 QY 121 KSMGNGKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158  
 DB 121 KSMGNGKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158

RESULT 2  
 US-08-468-413-2  
 ; Sequence 2, Application US/08468413  
 ; Patent No. 5861494  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LI, ET AL.  
 ; TITLE OF INVENTION: Human Colon Specific Gene  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
 ; STREET: 6 BECKER FARM ROAD  
 ; CITY: ROSELAND  
 ; STATE: NEW JERSEY  
 ; COUNTRY: USA  
 ; ZIP: 07068  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 INCH DISKETTE  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: WORD PERFECT 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/468,413  
 ; FILING DATE: 06 JUN 95  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: FERRARO, GREGORY D.  
 ; REGISTRATION NUMBER: 36,134  
 ; REFERENCE/DOCKET NUMBER: 325800-447  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 201-994-1700  
 ; TELEFAX: 201-994-1744  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 158 AMINO ACIDS  
 ; TYPE: AMINO ACID  
 ; STRANDEDNESS:  
 ; TOPOLOGY: LINEAR  
 ; MOLECULE TYPE: PROTEIN  
 ; US-08-468-413-2

Query Match 100.0%; Score 878; DB 2; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-88;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASRSMRLLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60  
 DB 1 MASRSMRLLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60  
 QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQWIDGAMLYYRSWSG 120  
 DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQWIDGAMLYYRSWSG 120  
 QY 121 KSMGNGKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158  
 DB 121 KSMGNGKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158

RESULT 3  
 US-09-162-508-2  
 ; Sequence 2, Application US/09162508  
 ; Patent No. 6080722  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LI, ET AL.  
 ; TITLE OF INVENTION: Human Colon Specific Gene  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
 ; STREET: 6 BECKER FARM ROAD  
 ; CITY: ROSELAND  
 ; STATE: NEW JERSEY  
 ; COUNTRY: USA  
 ; ZIP: 07068  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 INCH DISKETTE  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: WORD PERFECT 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/162,508  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/468,413  
 ; FILING DATE: 06 JUN 95  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: FERRARO, GREGORY D.  
 ; REGISTRATION NUMBER: 36,134  
 ; REFERENCE/DOCKET NUMBER: 325800-447  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 201-994-1700  
 ; TELEFAX: 201-994-1744  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 158 AMINO ACIDS  
 ; TYPE: AMINO ACID  
 ; STRANDEDNESS:  
 ; TOPOLOGY: LINEAR  
 ; MOLECULE TYPE: PROTEIN  
 ; US-09-162-508-2

Query Match 100.0%; Score 878; DB 3; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-88;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASRSMRLLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60  
 DB 1 MASRSMRLLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60  
 QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQWIDGAMLYYRSWSG 120  
 DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQWIDGAMLYYRSWSG 120  
 QY 121 KSMGNGKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158  
 DB 121 KSMGNGKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158

RESULT 4  
 PCT-US95-07169-2  
 ; Sequence 2, Application PC/TUS9507169  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LI, ET AL.  
 ; TITLE OF INVENTION: Human Colon Specific Gene  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
 ; STREET: 6 BECKER FARM ROAD



```

; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07169
; FILING DATE: 06 JUN 95
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-389
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; PCT-US95-07169-2
;
; Query Match 100.0%; Score 878; DB 5; Length 158;
; Best Local Similarity 100.0%; Pred. No. 5.9e-88;
; Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0
;
QY 1 MASRSRLRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECQS 60
Db 1 MASRSRLRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECQS 60
QY 61 YNGCAHLASTLSLKEASTIAEYISGYQRSOPITWGLHDHPQKRWQWIDGAMLYRWSWG 120
Db 61 YNGCAHLASTLSLKEASTIAEYISGYQRSOPITWGLHDHPQKRWQWIDGAMLYRWSWG 120
QY 121 KSMGKNKHCAEMSSNNFLTWSSNECKNRQHFCLKYRP 158
Db 121 KSMGKNKHCAEMSSNNFLTWSSNECKNRQHFCLKYRP 158
;
RESULT 5
US-09-621-976-4812
; Sequence 4812, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4812
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4812
;
; Query Match 79.2%; Score 695; DB 4; Length 122;
; Best Local Similarity 100.0%; Pred. No. 3.9e-69;
; Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

QY	26	MRPSCAPGWFYHKSNCYGYFRKLNRNSDALEECOSYNGNGAHLASILSLKEASTIAEYISG	85
Db	1	MRPSCAPGWFYHKSNCYGYFRKLNRNSDALEECOSYNGNGAHLASILSLKEASTIAEYISG	60
QY	86	YORSQPIWIGLHDPQKROQWIDGAMLYRWSGKSMGGNKHCAEMSSNNNFLTWSNSE	145
Db	61	YORSQPIWIGLHDPQKROQWIDGAMLYRWSGKSMGGNKHCAEMSSNNNFLTWSNSE	120
QY	146	CN 147	
Db	121	CN 122	
RESULT 6			
US-08-401-530A-7			
; Sequence 7, Application US/08401530A			
; Patent No. 5834590			
; GENERAL INFORMATION:			
; APPLICANT: Vinik, Aaron I.			
; APPLICANT: Pittenger, Gary L.			
; APPLICANT: Rafaeloff, Ronit			
; APPLICANT: Rosenber, Lawrence			
; APPLICANT: Duguid, William P.			
; TITLE OF INVENTION: INGA P PROTEIN INVOLVED IN PANCREATIC			
; TITLE OF INVENTION: ISLET NEOGENESIS			
; NUMBER OF SEQUENCES: 7			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Banner & Allegretti			
; STREET: 1001 G Street, N.W.			
; CITY: Washington			
; STATE: D.C.			
; COUNTRY: U.S.			
; ZIP: 20001-4597			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/401,530A			
; FILING DATE: 22-FEB-1995			
; CLASSIFICATION: 800			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Kagan, Sarah A.			
; REGISTRATION NUMBER: 32,141			
; REFERENCE/DOCKET NUMBER: 00570.48743			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 202-508-9100			
; TELEFAX: 202-508-9299			
; INFORMATION FOR SEQ ID NO: 7:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 165 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
; ORIGINAL SOURCE:			
; ORGANISM: Rattus rattus			
; US-08-401-530A-7			
Query Match 29.0%; Score 254.5; DB 2; Length 165;			
Best Local Similarity 30.6%; Pred. No. 6.6e-20;			
Matches 53; Conservative 38; Mismatches 51; Indels 29; Gaps 7			
QY	7	RLLLLSCLAKTGVGLDIIRPSS-----CAPGWFYHKSNCYGYFRKLNRW	51
Db	5	KYFILLSCL-----MVLSPSQGQAEEPLPSARITCPGGSNAYSSCYFFMEDHLWS	56
QY	52	SDAELECOSYNGNGAHLASILSLKEASTIAEYI--SGYORSQPIWIGLHDPKROQWID	109
Db	57	AEDALFQCNWNSG-YLVSVLSQAEGNFLASLIKESGTTAAN-VWIGLHDPKNNRRWHWS	114
QY	110	GAMLYRWSW-SGKSMWGNK-HCAEMSSNNNFLTWSNECNKRQHFCLKYR	157

DB 115 GSLFLYKSWDTGYPNNNRGVCVTSNSGYKKWRDNSCDAQLSFVCKFK 164

## RESULT 7

US-08-729-103-3  
; Sequence 3, Application US/08729103  
; Patent No. 5837841  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/729,103  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:

ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: P-0138 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 165 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 393209

US-08-729-103-3

Query Match 29.0%; Score 254.5; DB 2; Length 165;  
Best Local Similarity 30.6%; Pred. No. 6.6e-20;  
Matches 52; Conservative 38; Mismatches 51; Indels 29; Gaps 7;

QY	7	RLLLLSCLAKTGVLDIIMRPS-----CAPGWFYHKSNICYGFRKLRNW	51
DB	5	KYFILLSCL-----MVLSPSQQAEDLP SARITCPGSGNAYSSCYFMEHLWS	56
QY	52	SDAELECSQYGNAGHLASILSLKEASTIAEYI--SGYQSQPIWGLHDPKQKQWQWID	109
DB	57	AEADLFQNNMSG-YLVSVLSQAGNFLASLIKESGTTAAN-VWIGLHDPKNNRRHWSS	114
QY	110	GAMLYRSW-SGKSMGKNK-HCAEMSSNNFLTWSSNECNKQHFELCKYR	157
DB	115	GSLFLYKSWDTGYPNNNRGVCVTSNSGYKKWRDNSCDAQLSFVCKFK	164

## RESULT 8

US-08-709-662-7  
; Sequence 7, Application US/08709662  
; Patent No. 5840531  
; GENERAL INFORMATION:  
; APPLICANT: Winik, Aaron I.  
; APPLICANT: Pittenger, Gary L.

APPLICANT: Rafaeloff, Ronit  
APPLICANT: Rosenberg, Lawrence  
APPLICANT: Duguid, William P.  
TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC  
ISLET NEOGENESIS  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: US  
ZIP: 20001-4597

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/709,662  
; FILING DATE: 09-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 00570.59178  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 165 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Rattus rattus  
; US-08-709-662-7

Query Match 29.0%; Score 254.5; DB 2; Length 165;  
Best Local Similarity 30.6%; Pred. No. 6.6e-20;  
Matches 52; Conservative 38; Mismatches 51; Indels 29; Gaps 7;

QY	7	RLLLLSCLAKTGVLDIIMRPS-----CAPGWFYHKSNICYGFRKLRNW	51
DB	5	KYFILLSCL-----MVLSPSQQAEDLP SARITCPGSGNAYSSCYFMEHLWS	56
QY	52	SDAELECSQYGNAGHLASILSLKEASTIAEYI--SGYQSQPIWGLHDPKQKQWQWID	109
DB	57	AEADLFQNNMSG-YLVSVLSQAGNFLASLIKESGTTAAN-VWIGLHDPKNNRRHWSS	114
QY	110	GAMLYRSW-SGKSMGKNK-HCAEMSSNNFLTWSSNECNKQHFELCKYR	157
DB	115	GSLFLYKSWDTGYPNNNRGVCVTSNSGYKKWRDNSCDAQLSFVCKFK	164

## RESULT 9

US-08-729-103-4  
; Sequence 4, Application US/08729103  
; Patent No. 5837841  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,103
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0138 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 474306
; US-08-729-103-4

Query Match 27.3%; Score 240; DB 2; Length 166;
Best Local Similarity 34.3%; Pred. No. 2.6e-18;
Matches 46; Conservative 27; Mismatches 57; Indels 4; Gaps 4;

Qy 27 RPSCAPGWFYHKNCYGFPRKLRNWSDAELECQSYGNAGHLASTLSLKEASTIAEYISGY 86
Db 33 RISCPETNAYRSYCYFNEDEPETHVDADLYCQNMNSG-NLVSVLTQAEAGFVASLIKES 91
Qy 87 QRSQP-IWIGLHDPKQKQOWIDGAMLYRSW-SGKSMGNGK-HCAEMSSNNNFLTWS 143
Db 92 STDDSNVWIGLHDPKKNRRWHWSSGSLVSYKSWDTGSPSSANAGYCASLTSCSGFKKWKD 151
Qy 144 NECNKRQHFPLCKYR 157
Db 152 ESCEKFPFVCKFK 165

RESULT 10
US-09-949-016-10686
; Sequence 10686, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10686
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-10686

Query Match 27.3%; Score 240; DB 4; Length 174;
Best Local Similarity 34.3%; Pred. No. 2.7e-18;

```

```

Matches 46; Conservative 27; Mismatches 57; Indels 4; Gaps 4;

Qy 27 RPSCAPGWFYHKNCYGFPRKLRNWSDAELECQSYGNAGHLASTLSLKEASTIAEYISGY 86
Db 41 RISCPETNAYRSYCYFNEDEPETHVDADLYCQNMNSG-NLVSVLTQAEAGFVASLIKES 99
Qy 87 QRSQP-IWIGLHDPKQKQOWIDGAMLYRSW-SGKSMGNGK-HCAEMSSNNNFLTWS 143
Db 100 STDDSNVWIGLHDPKKNRRWHWSSGSLVSYKSWDTGSPSSANAGYCASLTSCSGFKKWKD 159
Qy 144 NECNKRQHFPLCKYR 157
Db 160 ESCEKFPFVCKFK 173

RESULT 11
US-08-401-530A-5
; Sequence 5, Application US/08401530A
; Patent No. 5834590
; GENERAL INFORMATION:
; APPLICANT: Vinik, Aaron I.
; APPLICANT: Pittenger, Gary L.
; APPLICANT: Rafaeloff, Ronit
; APPLICANT: Rosenberg, Lawrence
; APPLICANT: Duguid, William F.
; TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC
; ISLET NEOGENESIS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/401,530A
; FILING DATE: 22-FEB-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 00570.48743
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Rattus rattus
; US-08-401-530A-5

Query Match 26.3%; Score 231; DB 2; Length 174;
Best Local Similarity 32.4%; Pred. No. 2.6e-17;
Matches 55; Conservative 28; Mismatches 73; Indels 14; Gaps 5;

Qy 1 MASRSMLLLLSCLAKTGVLD-----IIMPSCAPGWFYHKNCYGFPRKLRNWSDA 54
Db 5 VALTMSWMLLSLMLLSQVGEDAKEDVPTSRISCPKSRAYGSCYALFVSFVSKSWFDA 64
Qy 55 ELECQSYGNAGHLASTLSLKEASTIAEYI-SGYQRSQPIWIGLHDPKQKQO-----WQVI 108
Db 65 DLACQKRPSPG-HLVSVLVSSEASFVSLIKSSGNSGQNVWIGLHDPKQKQPNRGWWS 123

```

QY 109 DGAMVLYRSW-SGKSMGNGKHCAMSSNNFLTWSNECKRQHFLCKYR 157  
Db 124 NADVMYFNWETNPSSVSGHCGTLTRASGFLRWRENNCISELPYVCKFK 173

## RESULT 12

US-08-709-662-5  
; Sequence 5, Application US/08709662  
; Patent No. 5840531  
; GENERAL INFORMATION:  
; APPLICANT: Vink, Aaron I.  
; APPLICANT: Pittenger, Gary L.  
; APPLICANT: Rafaeloff, Ronit  
; APPLICANT: Rosenberg, Lawrence  
; APPLICANT: Duguid, William P.  
; TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: US  
; ZIP: 20001-4597  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08709,662  
; APPLICATION NUMBER: 09-SEP-1996  
; FILING DATE: 09-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 00570.59178  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 174 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Rattus rattus  
; US-08-709-662-5

Query Match 26.3%; Score 231; DB 2; Length 174;  
Best Local Similarity 32.4%; Pred. No. 2.6e-17;  
Matches 55; Conservative 28; Mismatches 73; Indels 14; Gaps 5;

QY 1 MASRSMLLLLSCLAKTGVLD-----IMRPSCAPGWFYHKSNICYGYFKLRNWSDA 54  
Db 5 VALTTSMWLLSLLSLQVQGEKEDVDPTSRISCPKGRAYGSCYALFVSWSKWFDA 64  
QY 55 ELECCSYGNGAHLASILSKKEASTIAEYI-SGYORSQPIWIGLHDPQKRQ-----WQMI 108  
Db 65 DLACQKRPSPG-HLVSVLSGSEAFVSSLIKSGNSGQNVWIGLHDPQLGQENRGGEWS 123  
QY 109 DGAMVLYRSW-SGKSMGNGKHCAMSSNNFLTWSNECKRQHFLCKYR 157  
Db 124 NADVMYFNWETNPSSVSGHCGTLTRASGFLRWRENNCISELPYVCKFK 173

## RESULT 13

US-08-401-530A-6  
; Sequence 6, Application US/08401530A  
; Patent No. 5834590  
; GENERAL INFORMATION:

; APPLICANT: Vink, Aaron I.  
; APPLICANT: Pittenger, Gary L.  
; APPLICANT: Rafaeloff, Ronit  
; APPLICANT: Rosenberg, Lawrence  
; APPLICANT: Duguid, William P.  
; TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Allegretti  
; STREET: 1001 G Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: US  
; ZIP: 20001-4597  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/401,530A  
; APPLICATION NUMBER: 22-FEB-1995  
; FILING DATE: 22-FEB-1995  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 00570.48743  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 174 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Rattus rattus  
; US-08-401-530A-6

Query Match 26.0%; Score 228.5; DB 2; Length 174;  
Best Local Similarity 34.5%; Pred. No. 4.9e-17;  
Matches 48; Conservative 25; Mismatches 57; Indels 9; Gaps 4;  
QY 27 RPSCAPGWFYHKSNICYGYFKLRNWSDAELECCSYGNGAHLASILSKKEASTIAEYISG- 85  
Db 36 RTSCPMGSKAYRSYCYTLVTTLKSWFQADLACQKRPSPG-HLVSVLSGSEAFVSSLIWTR 94  
QY 86 YORSQPIWIGLHDPQKRQ-----WQIDCAMVLYRSWSG--KSMGNGKHCAMSSNNNF 138  
Db 95 VNNQDITWILHDPMTGQPNQGGGWSNSDVLYNLNWDGDPSSTVNRGNGSLTATSEF 154  
QY 139 LTWSSNECKRQHFLCKYR 157  
Db 155 LKWGDHCHCDVLPFVCKFK 173

## RESULT 14

US-08-709-662-6  
; Sequence 6, Application US/08709662  
; Patent No. 5840531  
; GENERAL INFORMATION:

; APPLICANT: Vink, Aaron I.  
; APPLICANT: Pittenger, Gary L.  
; APPLICANT: Rafaeloff, Ronit  
; APPLICANT: Rosenberg, Lawrence  
; APPLICANT: Duguid, William P.  
; TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.

STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: US  
ZIP: 20001-4597  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/709,662  
FILING DATE: 09-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 00570.59178  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Rattus rattus  
US-08-709-662-6

Query Match	26.0%;	Score 228.5;	DB 2;	Length 174;
Best Local Similarity	34.5%;	Pred. No. 4.9e-17;		
Matches	48;	Conservative 25;	Mismatches 57;	Indels 9; Gaps 4;
Qy	27	RPSCAPGFVHKSNCYGFRKLRNWSDAELECSQVGNCAHLASTLSLKEASTIAEYISG- 85		
Db	36	RTSCPMSGKAYSICYTLVLTLSKWFQADLACQKRPSG-HLVSLTSGEASFVSSLVTR 94		
Qy	86	YQRSQPTIWIGLHPQKRQQ-----WQIDGAMLYLYRSWG--KSMGGNGKHCAMSSNNNF 138		
Db	95	VNNNQDIWIWLHPTWTGQQPNGGGMWNSDVLNLNWDGDPSSVTVRNGCGSLTATSEF 154		
Qy	139	LTWSSNECNKRQHFLLCKYR 157		
Db	155	LKQGDHHCDDVELPFCVKFK 173		

RESULT 15

```

RES001 13
US-09-949-016-10685
; Sequence 10685, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE OF INVENTION: WITH HUMAN DISEASE,
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10685
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10685

```

```

Query Match      25.5%; Score 224; DB 4; Length 144;
Best Local Similarity 33.3%; Pred. No. 1.2e-16;
Matches 45; Conservative 26; Mismatches 58; Indels 6; Gaps 4

Qy 27 RPSCAPGFYHKSCYCYGFRKLRNWSDAELSCQYGVNGAHLASILSLKEASTIAEYI--S 84
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 RISCEPGETNAYRSCYYFENEDRETWTWDADLYCQNMNSG-NLVSVLTAQEGAFVASLIKES 69
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 85 GYRSDPIWGLHDPKQKQWHDGAMLYRSW--SGKSMGNGKXKCAEMSGNNNFLTWS 142
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 GTDDFN-VWTLGLHDPKKNRRHWHSSGSLVSYKSGWGAPSVNPGYCVSLTSTSTGFQKWK 128
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 143 SNECNKQHFLLCKYR 157
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 DVPCEDKFSFVCKEK 143
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: March 3, 2005, 07:55:36
Job time : 23 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 07:54:33 ; Search time 132 Seconds  
(without alignments)  
393.857 Million cell updates/sec

Title: US-10-099-791E-2

Perfect score: 878  
Sequence: 1 MASRMRLLLLLSCLAKTV.....LTWSSNECKRQHFLCKYRP 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 32904822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878	100.0	158	9	US-09-922-217-1070 Sequence 1070, Ap
2	878	100.0	158	9	US-09-922-217-1077 Sequence 1077, Ap
3	878	100.0	158	9	US-09-922-217-1078 Sequence 1078, Ap
4	878	100.0	158	9	US-09-922-217-1079 Sequence 1079, Ap
5	878	100.0	158	9	US-09-922-217-1080 Sequence 1080, Ap
6	878	100.0	158	9	US-09-833-263-1070 Sequence 1070, Ap
7	878	100.0	158	9	US-09-833-263-1077 Sequence 1077, Ap
8	878	100.0	158	9	US-09-833-263-1078 Sequence 1078, Ap
9	878	100.0	158	9	US-09-833-263-1079 Sequence 1079, Ap
10	878	100.0	158	9	US-09-833-263-1080 Sequence 1080, Ap
11	878	100.0	158	10	US-09-525-041-2 Sequence 2, Appli
12	878	100.0	158	11	US-09-969-034-4471 Sequence 4471, Ap
13	878	100.0	158	13	US-10-025-380-1070 Sequence 1070, Ap

14	878	100.0	158	13	US-10-025-380-1077 Sequence 1077, Ap
15	878	100.0	158	13	US-10-025-380-1078 Sequence 1078, Ap
16	878	100.0	158	13	US-10-025-380-1079 Sequence 1079, Ap
17	878	100.0	158	13	US-10-025-380-1080 Sequence 1080, Ap
18	878	100.0	158	14	US-10-100-608B-2 Sequence 2, Appli
19	878	100.0	158	14	US-10-157-031-114 Sequence 114, App
20	878	100.0	158	14	US-10-205-823-347 Sequence 347, App
21	878	100.0	158	15	US-10-295-027-138 Sequence 138, App
22	878	100.0	158	15	US-10-295-027-781 Sequence 781, App
23	878	100.0	158	15	US-10-295-027-861 Sequence 861, App
24	878	100.0	158	15	US-10-173-999-93 Sequence 93, Appl
25	878	100.0	158	15	US-10-173-999-93 Sequence 93, Appl
26	878	100.0	158	16	US-10-734-564-105 Sequence 105, App
27	878	100.0	158	16	US-10-099-791E-2 Sequence 2, Appli
28	878	100.0	166	14	US-10-106-698-6394 Sequence 9, Appli
29	760	86.6	367	15	US-10-452-646-9 Sequence 753, Appl
30	240	27.3	174	9	US-09-925-297-753 Sequence 113, App
31	231	26.3	174	15	US-10-028-248A-113 Sequence 113, App
32	231	26.3	174	15	US-10-107-782-113 Sequence 110, App
33	231	26.3	175	15	US-10-028-248A-110 Sequence 110, App
34	231	26.3	175	15	US-10-107-782-110 Sequence 30, Appl
35	225.5	25.7	166	10	US-09-997-003-30 Sequence 43, Appl
36	225.5	25.7	166	10	US-09-997-003-43 Sequence 4, Appli
37	225.5	25.7	166	16	US-10-734-564-4 Sequence 1182, Ap
38	225.5	25.7	174	9	US-09-925-301-1182 Sequence 2, Appli
39	224	25.5	166	16	US-10-734-564-2 Sequence 38, Appl
40	223	25.4	175	15	US-10-028-248A-38 Sequence 38, Appl
41	223	25.4	175	15	US-10-107-782-38 Sequence 3, Appli
42	221	25.2	175	14	US-10-316-761-3 Sequence 1, Appli
43	221	25.2	175	15	US-10-434-906-1 Sequence 109, App
44	221	25.2	175	15	US-10-107-782-109 Sequence 109, App
45	221	25.2	175	15	US-10-107-782-109 Sequence 109, App

ALIGNMENTS

RESULT 1

US-09-922-217-1070  
; Sequence 1070, Application US/09922217  
; Patent No. US2002007641A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Smith, Carole Lynn  
; APPLICANT: Wang, Gordon E.  
; APPLICANT: Clapper, Jonathan D.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.471C13  
; CURRENT APPLICATION NUMBER: US/09/922.217  
; CURRENT FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 1124  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1070  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-922-217-1070

Query Match 100.0%; Score 878; DB 9; Length 158;  
Best Local Similarity 100.0%; Pred. No. 3.3e-82;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASRMRLLLLLSCLAKTVGLDILMPSCAPGWFYKSNICYGFRKLRNWSDELECCS 60

Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECS 60  
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPTIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPTIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
QY 121 KSMGNKHCAMSSNNFLTWSNECKRQHFLCKYRP 158  
Db 121 KSMGNKHCAMSSNNFLTWSNECKRQHFLCKYRP 158

## RESULT 2

US-09-922-217-1077  
; Sequence 1077, Application US/09922217  
; Patent No. US20020076414A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Smith, Carole Lynn  
; APPLICANT: King, Gordon E.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; FILE REFERENCE: 210121.471C13  
; CURRENT APPLICATION NUMBER: US/09/922,217  
; NUMBER OF SEQ ID NOS: 1124  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1077  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-922-217-1077

Query Match 100.0%; Score 878; DB 9; Length 158;  
Best Local Similarity 100.0%; Pred. No. 3.3e-82;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECS 60  
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECS 60  
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPTIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPTIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
QY 121 KSMGNKHCAMSSNNFLTWSNECKRQHFLCKYRP 158  
Db 121 KSMGNKHCAMSSNNFLTWSNECKRQHFLCKYRP 158

## RESULT 3

US-09-922-217-1078  
; Sequence 1078, Application US/09922217  
; Patent No. US20020076414A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Smith, Carole Lynn  
; APPLICANT: King, Gordon E.

; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; FILE REFERENCE: 210121.471C13  
; CURRENT APPLICATION NUMBER: US/09/922,217  
; CURRENT FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 1124  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1078  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-922-217-1078

Query Match 100.0%; Score 878; DB 9; Length 158;  
Best Local Similarity 100.0%; Pred. No. 3.3e-82;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECS 60  
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECS 60  
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPTIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPTIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
QY 121 KSMGNKHCAMSSNNFLTWSNECKRQHFLCKYRP 158  
Db 121 KSMGNKHCAMSSNNFLTWSNECKRQHFLCKYRP 158

## RESULT 4

US-09-922-217-1079  
; Sequence 1079, Application US/09922217  
; Patent No. US20020076414A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Smith, Carole Lynn  
; APPLICANT: King, Gordon E.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; FILE REFERENCE: 210121.471C13  
; CURRENT APPLICATION NUMBER: US/09/922,217  
; CURRENT FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 1124  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1079  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-922-217-1079

Query Match 100.0%; Score 878; DB 9; Length 158;  
Best Local Similarity 100.0%; Pred. No. 3.3e-82;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECS 60  
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECS 60  
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPTIWIGLHDPKQKQOWIDGAMLYRSWSG 120  
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPTIWIGLHDPKQKQOWIDGAMLYRSWSG 120



QY 121 KSMGKNGKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158  
 Db 121 KSMGKNGKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 5

US-09-922-217-1080  
 ; Sequence 1080, Application US/09922217  
 ; Patent No. US2002007641A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jianchun  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Secrist, Heather  
 ; APPLICANT: Benson, Darin R.  
 ; APPLICANT: Meagher, Madeleine Joy  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Smith, Carole Lynn  
 ; APPLICANT: King, Gordon E.  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Clapper, Jonathan D.  
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
 ; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 210121.471C13  
 ; CURRENT APPLICATION NUMBER: US/09/922.217  
 ; CURRENT FILING DATE: 2001-08-03  
 ; NUMBER OF SEQ ID NOS: 1124  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1080  
 ; LENGTH: 158  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-922-217-1080

Query Match 100.0%; Score 878; DB 9; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-82;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNKCYGYFRKLRNWSDALEECQS 60  
 Db 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNKCYGYFRKLRNWSDALEECQS 60  
 QY 61 YNGAHLASLILSLKEASTIAEYISGYORSQPIWGLHDPKQROQWIDGAMLYRSWSG 120  
 Db 61 YNGAHLASLILSLKEASTIAEYISGYORSQPIWGLHDPKQROQWIDGAMLYRSWSG 120  
 QY 121 KSMGKNGKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158  
 Db 121 KSMGKNGKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 6

US-09-833-263-1070  
 ; Sequence 1070, Application US/09833263  
 ; Patent No. US20020110547A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Clapper, Jonathan D.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Meagher, Madeleine J.  
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 210121.471C12  
 ; CURRENT APPLICATION NUMBER: US/09/833.263  
 ; CURRENT FILING DATE: 2001-04-10  
 ; NUMBER OF SEQ ID NOS: 1093  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1070  
 ; LENGTH: 158  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

US-09-833-263-1070

Query Match 100.0%; Score 878; DB 9; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-82;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNKCYGYFRKLRNWSDALEECQS 60  
 Db 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNKCYGYFRKLRNWSDALEECQS 60  
 QY 61 YNGAHLASLILSLKEASTIAEYISGYORSQPIWGLHDPKQROQWIDGAMLYRSWSG 120  
 Db 61 YNGAHLASLILSLKEASTIAEYISGYORSQPIWGLHDPKQROQWIDGAMLYRSWSG 120  
 QY 121 KSMGKNGKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158  
 Db 121 KSMGKNGKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 7

US-09-833-263-1077  
 ; Sequence 1077, Application US/09833263  
 ; Patent No. US20020110547A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Clapper, Jonathan D.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Meagher, Madeleine J.  
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 210121.471C12  
 ; CURRENT APPLICATION NUMBER: US/09/833.263  
 ; CURRENT FILING DATE: 2001-04-10  
 ; NUMBER OF SEQ ID NOS: 1093  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1077  
 ; LENGTH: 158  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-833-263-1077

Query Match 100.0%; Score 878; DB 9; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-82;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNKCYGYFRKLRNWSDALEECQS 60  
 Db 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNKCYGYFRKLRNWSDALEECQS 60  
 QY 61 YNGAHLASLILSLKEASTIAEYISGYORSQPIWGLHDPKQROQWIDGAMLYRSWSG 120  
 Db 61 YNGAHLASLILSLKEASTIAEYISGYORSQPIWGLHDPKQROQWIDGAMLYRSWSG 120  
 QY 121 KSMGKNGKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158  
 Db 121 KSMGKNGKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 8

US-09-833-263-1078  
 ; Sequence 1078, Application US/09833263  
 ; Patent No. US20020110547A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Clapper, Jonathan D.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Meagher, Madeleine J.  
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 210121.471C12  
 ; CURRENT APPLICATION NUMBER: US/09/833.263  
 ; CURRENT FILING DATE: 2001-04-10  
 ; NUMBER OF SEQ ID NOS: 1093

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1078
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1078

Query Match      100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMLLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECOQ 60
    |||||||
Db 1 MASRSMLLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECOQ 60
    |||||||
QY 61 YNGGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
    |||||||
Db 61 YNGGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
    |||||||
QY 121 KSMGKNKHCHEMSSNNNFLTWSNNECKRQHFLCKYRP 158
    |||||||
Db 121 KSMGKNKHCHEMSSNNNFLTWSNNECKRQHFLCKYRP 158
    |||||||

RESULT 9
US-09-833-263-1079
; Sequence 1079, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stoik, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1079
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1079

Query Match      100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMLLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECOQ 60
    |||||||
Db 1 MASRSMLLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECOQ 60
    |||||||
QY 61 YNGGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
    |||||||
Db 61 YNGGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
    |||||||
QY 121 KSMGKNKHCHEMSSNNNFLTWSNNECKRQHFLCKYRP 158
    |||||||
Db 121 KSMGKNKHCHEMSSNNNFLTWSNNECKRQHFLCKYRP 158
    |||||||

RESULT 10
US-09-833-263-1080
; Sequence 1080, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stoik, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
```

```
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1080
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1080

Query Match      100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMLLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECOQ 60
    |||||||
Db 1 MASRSMLLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECOQ 60
    |||||||
QY 61 YNGGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
    |||||||
Db 61 YNGGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
    |||||||
QY 121 KSMGKNKHCHEMSSNNNFLTWSNNECKRQHFLCKYRP 158
    |||||||
Db 121 KSMGKNKHCHEMSSNNNFLTWSNNECKRQHFLCKYRP 158
    |||||||

RESULT 11
US-09-525-041-2
; Sequence 2, Application US/09525041
; Publication No. US20030158098A1
; GENERAL INFORMATION:
; APPLICANT: Soppet et al.
; TITLE OF INVENTION: Colon Specific Gene and Protein
; FILE REFERENCE: PF178D2
; CURRENT APPLICATION NUMBER: US/09/525,041
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 09/162,508
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: US 08/468,413
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-525-041-2

Query Match      100.0%; Score 878; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMLLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECOQ 60
    |||||||
Db 1 MASRSMLLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECOQ 60
    |||||||
QY 61 YNGGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
    |||||||
Db 61 YNGGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
    |||||||
QY 121 KSMGKNKHCHEMSSNNNFLTWSNNECKRQHFLCKYRP 158
    |||||||
Db 121 KSMGKNKHCHEMSSNNNFLTWSNNECKRQHFLCKYRP 158
    |||||||

RESULT 12
US-09-969-034-4471
; Sequence 4471, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4471
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-034-4471

Query Match      100.0%; Score 878; DB 11; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60

QY 61 YNGAHLASLSLKEASTIAEYISGYQSORQPIWIGLHDPKQROQWIDGAMLYRSWSG 120
DB 61 YNGAHLASLSLKEASTIAEYISGYQSORQPIWIGLHDPKQROQWIDGAMLYRSWSG 120

QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158
DB 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158

```

```

RESULT 13
US-10-025-380-1070
; Sequence 1070, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1070
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1070

```

```

Query Match      100.0%; Score 878; DB 13; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60

QY 61 YNGAHLASLSLKEASTIAEYISGYQSORQPIWIGLHDPKQROQWIDGAMLYRSWSG 120
DB 61 YNGAHLASLSLKEASTIAEYISGYQSORQPIWIGLHDPKQROQWIDGAMLYRSWSG 120

QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158
DB 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158

```

```

RESULT 15
US-10-025-380-1078
; Sequence 1078, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1077
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1077

```

```

; GENERAL INFORMATION:
; APPLICANT: Xu Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1078
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1078

```

```

Query Match      100.0%; Score 878; DB 13; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-82;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MASRSMRLLLLSCLAKTGVLGDIIMRPSCAPGFYHKSNICYGFRKLRNWSDALEQCS 60
Db      1 MASRSMRLLLLSCLAKTGVLGDIIMRPSCAPGFYHKSNICYGFRKLRNWSDALEQCS 60

Qy      61 YGNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPKROQOWIDGAMLYRSWSG 120
Db      61 YGNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPKROQOWIDGAMLYRSWSG 120

Qy      121 KSMGNGKHCAMSSNNFLTWSSNECNKQHPCKYRP 158
Db      121 KSMGNGKHCAMSSNNFLTWSSNECNKQHPCKYRP 158

```

Search completed: March 3, 2005, 08:06:40  
Job time : 133 secs





Db 92 STDDSNVIGLHDPKQNRWHSWSSLSVYSKSWDTGSPSSANAGYCASLTSCSGFKKWKD 151

Qy 144 NECNKRQHFLCKYR 157  
| | | | |  
| | | | |

Db 152 ESECKKFSVCKFK 165  
| | | | |  
| | | | |

RESULT 5

B47148  
reg II, regenerating islet cells - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C:Accession: B47148

R;Unno, M.; Yonekura, H.; Nakagawara, K.; Watanabe, T.; Miyashita, H.; Morizumi, S.; Ok  
J. Biol. Chem. 268, 15974-15982, 1993

A:Title: Structure, chromosomal localization, and expression of mouse reg genes, reg I a  
A:Reference number: A47148; MUID: 93340209; PMID: 8340418

A:Accession: B47148

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-173 <UNN>

A:Cross-references: UNIPROT:Q08731; GB:D14011; NID:g391773; PIDN:BAA03112.1; PID:g391774

C:Genetics:

A:Introns: 22/1; 68/3; 114/3; 152/1

C:Superfamily: tetranectin; C-type lectin homology

F:43-169/Domain: C-type lectin homology <LCH>

F:54-71-169,144-161/Disulfide bonds: #status predicted

Query Match 26.9%; Score 236.5; DB 2; Length 173;  
Best Local Similarity 28.7%; Pred. No. 5.8e-16;  
Matches 50; Conservative 33; Mismatches 72; Indels 19; Gaps 5;

Qy 1 MASRMRLLLLSLAKTGVLDII-----MRPSCAPGWYFKSNVCYFPRK 47  
| | | | |  
| | | | |

Db 1 MAQNNVYILFLCLMFLSYSGQVAEDFPLEAKDLPSAKINCPGANAYSCYCYLIED 60  
| | | | |  
| | | | |

Qy 48 LRNWSDALEQCSYNGNHLASILSKASTIAEYI--SGYQSQPIWIGLHDPKQKQW 105  
| | | | |  
| | | | |

Db 61 RLTWGEADLFQNM-NAGHLVSLQAESNFVASLVKESGTTASN-VMTGLHDPKSNRW 118  
| | | | |  
| | | | |

Qy 106 QWIDGMVLYFSMS--GKSMGNKHCAMSSNNFLTWSSNECKRQHFLCKYR 157  
| | | | |  
| | | | |

Db 119 HWSSGSUFLFKSWATGAPSTANRGYCVSLTNTAYKKWKOENCAQYSVCKFR 172  
| | | | |  
| | | | |

RESULT 6

183377  
regenerating protein III (reg III) - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
C:Accession: I60296; I83377

R;Suzuki, Y.; Yonekura, H.; Watanabe, T.; Unno, M.; Morizumi, S.; Miyashita, H.; Okamoto  
Gene 144, 315-316, 1994

A:Title: Structure and expression of a novel rat RegIII gene.  
A:Reference number: I60296; MUID: 94314238; PMID: 8039722

A:Accession: I60296

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-174 <RES>

A:Cross-references: UNIPROT:P35231; GB:D23676; NID:g471157; PIDN:BAA04904.1; PID:g471158

A:Accession: I83377

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-174 <RE2>

A:Cross-references: GB:D26078; NID:g471159; PIDN:BAA05071.1; PID:g471160

C:Genetics:

A:Gene: RegIII

A:Introns: 25/1; 64/3; 110/3; 153/1

C:Superfamily: tetranectin; C-type lectin homology

F:33-170/Domain: C-type lectin homology <LCH>

Query Match 26.9%; Score 236.5; DB 2; Length 174;  
Best Local Similarity 35.3%; Pred. No. 5.8e-16;

Matches 49; Conservative 25; Mismatches 56; Indels 9; Gaps 4;

Qy 27 RPSCAPGWYFKSNVCYFRLRNWSDAELEQCSYNGNHLASILSKASTIAEYISG- 85  
| | | | |  
| | | | |

Db 36 RTSCPMGSKAYRSYCYTLVTTLSKSWFQADLACQKRPSSG-HLVSLSGGEASFVSSLVTR 94  
| | | | |  
| | | | |

Qy 86 YQSQPIWIGLHDPKQKQ-----QWIDGMVLYFSMS--KSMGNKHCAMSSNNNF 138  
| | | | |  
| | | | |

Db 95 VNNQDIWIGLHDPKQKQNGGWSNSDVLNWLNDGDPSTVNRNGCSLTATSEF 154  
| | | | |  
| | | | |

Qy 139 LWTSSNECKRQHFLCKYR 157  
| | | | |  
| | | | |

Db 155 LKWDHHCDELFPVCKFK 173  
| | | | |  
| | | | |

RESULT 7

A38609  
lectin, galactose-specific - western diamondback rattlesnake  
C:Species: Crotalus atrox (western diamondback rattlesnake)  
C:Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 09-Jul-2004  
C:Accession: A38609

R;Hirabayashi, J.; Kusunoki, T.; Kasai, K.  
J. Biol. Chem. 266, 2320-2326, 1991

A:Title: Complete primary structure of a galactose-specific lectin from the venom of the  
A:Reference number: A38609; MUID: 91115849; PMID: 1989986

A:Accession: A38609

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-135 <HIR>

A:Cross-references: UNIPROT:P21963

C:Superfamily: tetranectin; C-type lectin homology

F:3-131/Domain: C-type lectin homology <LCH>

F:3-14,31-131,106-123/Disulfide bonds: #status predicted

Query Match 26.5%; Score 233; DB 2; Length 135;  
Best Local Similarity 34.4%; Pred. No. 9.8e-16;  
Matches 45; Conservative 18; Mismatches 64; Indels 4; Gaps 2;

Qy 29 SCAPGWYFKSNVCYFRLRNWSDAELEQCSYNGNHLASILSKASTIAEYISGYQR 88  
| | | | |  
| | | | |

Db 2 NCPLDLPNGLNLCYKIFNLKWTDEAEFCKYKPGCHLASFHYRGESLEIAEYISDYHK 61  
| | | | |  
| | | | |

Qy 89 SQP-IWIGLHDPKQKQWIDGMVLYFSWSGK---MGNGKHCAMSSNNFLTWSSN 144  
| | | | |  
| | | | |

Db 62 QGVNWIQLRDKKDFSWETDRCTDYLTDKNQPDHYQNKFCVELSLTGYRLWNDQ 121  
| | | | |  
| | | | |

Qy 145 ECNKRQHFLCK 155  
| | | | |  
| | | | |

Db 122 VCESKDAFLCQ 132  
| | | | |  
| | | | |

RESULT 8

S54979  
pancreatitis-associated protein PAP-3 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
C:Accession: S54979; S43438

R;Dusetti, N.J.; Frigerio, J.M.; Szpirer, C.; Dagorn, J.C.; Iovanna, J.L.  
Biochem. J. 307, 9-16, 1995

A:Title: Cloning, expression and chromosomal localization of the rat pancreatitis-associat  
A:Reference number: S54979; MUID: 95234061; PMID: 7717998

A:Accession: S54979

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-174 <DUS>

A:Cross-references: UNIPROT:P42854; EMBL:U09193; NID:g489391; PIDN:AAA79231.1; PID:g48939;  
R;Frigerio, J.M.; Dusetti, N.J.; Garrido, P.; Dagorn, J.C.; Iovanna, J.L.  
Biochim. Biophys. Acta 1216, 329-331, 1993

A:Title: The pancreatitis associated protein III (PAP III), a new member of the PAP gene  
A:Reference number: S43438; MUID: 94060113; PMID: 8241280

A:Accession: S43438

A:Status: preliminary

A:Molecule type: mRNA

A;Residues: 1-174 <FRI>  
A;Cross-references: GB:L20869; GB:S67496; NID:g463279; PIDN:AAA1809.1; PID:g463280  
C;Genetics:  
A;Introns: 26/1; 65/3; 111/3; 153/1  
A;Superfamily: tetranectin; C-type lectin homology  
F;40-170/Domain: C-type lectin homology <LCH>  
Query Match 26.3%; Score 231; DB 2; Length 174;  
Best Local Similarity 32.4%; Pred. No. 2e-15; Indels 14; Gaps 5;  
Matches 55; Conservative 28; Mismatches 73  
QY 1 MASRNRLLLLSLCLATKTVLGD-----IMRPSCAPGFYHKSNCGYFRKLRNWSDA 54  
DB 5 VALTITSMMLLSLMLLSLQVQSGDAKEDVPTSRISCPKGRAYSGSYCALFVSRSKSWFDA 64  
QY 55 ELECCSYGNGAHLASILSKEASTIAIYI-SCYORSQPIWGLHDPKQKQ-----WQWI 108  
DB 65 DLACQKRFSG-HLVSVLSGSEAFVSLSLKSSGNGQNVWVGLHDPGLQGPENRGGWWS 123  
QY 109 DGAMLYLRSG-SGKSMGNGKHCAEWSNNNFLTWSNNECNKROHFLCKYR 157  
DB 124 NADVMNYPNWTNPSSVSGSHCGTLTRASGFLRWRENNCISELPYVCKFK 173  
RESULT 9  
A37194  
pancreatic thread protein precursor - bovine  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 31-Jan-1992 #sequence\_revision 31-Jan-1992 #text\_change 09-Jul-2004  
C;Accession: A37194; A53897  
R;de la Monte, S.M.; Ozturk, M.; Wands, J.R.  
J. Clin. Invest. 86, 1004-1013, 1990  
A;Title: Enhanced expression of an exocrine pancreatic protein in Alzheimer's disease and in transgenic mice.  
A;Reference number: A37194; MUID:90368981; PMID:2394826  
A;Accession: A37194  
A;Molecule type: mRNA  
A;Residues: 1-175 <DEL>  
A;Cross-references: UNIPROT:P23132; GB:M59794; NID:g163648; PIDN:AAA30750.1; PID:g163649  
R;Cai, L.; Harris, W.R.; Marshak, D.R.; Gross, J.; Crabb, J.W.  
J. Protein Chem. 9, 623-632, 1990  
A;Title: Structural analysis of bovine pancreatic thread protein.  
A;Reference number: A53897; MUID:91197388; PMID:2085387  
A;Accession: A53897  
A;Molecule type: protein  
A;Residues: 38-138;141-175 <CAT>  
C;Comment: the purified protein undergoes a reversible globule-fibril transformation and is stable in the presence of urea.  
C;Superfamily: tetranectin; C-type lectin homology  
C;Keywords: disulfide bond; extracellular protein; pancreas  
F;38-138/Product: pancreatic thread protein chain A #status experimental <ACH>  
F;40-171/Domain: C-type lectin homology <LCH>  
F;141-175/Product: pancreatic thread protein chain B #status experimental <BCH>  
Query Match 26.3%; Score 231; DB 2; Length 175;  
Best Local Similarity 30.9%; Pred. No. 2e-15; Indels 22; Gaps 6;  
Matches 51; Conservative 30; Mismatches 62  
QY 10 LLLSLCLATKTVLGD-----IMRPSCAPGFYHKSNCGYFRKLRNWSDALECCQ 59  
DB 13 MLLSLCLM-----LLSQIQENSGKELPSARISCPGSMAYRSHCYALFXTPTKWNMDADIACQ 69  
QY 60 SYNGCAHLASILSKEASTIAIYISQYORSQ-PIWGLHDPKQKQ-----WQWIDGAMY 113  
DB 70 KRPSG-HLVSVLSGAEESFVASLVRNLTQSDIWGLHDPTEGSEANAGGWWSINDVL 128  
QY 114 LYRSG-SGKSMGNGKHCAEWSNNNFLTWSNNECNKROHFLCKY 156  
DB 129 NYVAMETDPAISIPFGYCGSLSRSSGYLKWDRHNCNLPVCKFK 173  
RESULT 10  
A48689  
pancreatitis-associated protein PAP-2 - rat  
C;Species: Rattus norvegicus (Norway rat)

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: A48689  
R;Frigerio, J.M.; Duseti, N.J.; Keim, V.; Dagorn, J.C.; Iovanna, J.L.  
Biochemistry 32, 9236-9241, 1993  
A;Title: Identification of a second rat pancreatitis-associated protein. Messenger RNA cDNA  
A;Reference number: A48689; MUID:93378971; PMID:8369291  
A;Accession: A48689  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-174 <FRI>  
A;Cross-references: UNIPROT:P35231; GB:L10229; NID:g409014; PIDN:AAA02980.1; PID:g409015  
C;Superfamily: tetranectin; C-type lectin homology  
F;39-170/Domain: C-type lectin homology <LCH>  
F;39-50;67-170;145-162/Disulfide bonds: #status predicted  
Query Match 26.0%; Score 228.5; DB 1; Length 174;  
Best Local Similarity 34.5%; Pred. No. 3.6e-15; Indels 9; Gaps 4;  
Matches 48; Conservative 25; Mismatches 57  
QY 27 RPSCAPGFYHKSNCGYFRKLRNWSDALECCSYGNGAHLASILSKEASTIAIYI 85  
DB 36 RTSCPMGSKAYRSYCYTLVTLKSWFOADLACQKRFSG-HLVSVLSGSEAFVSLSVTGR 94  
QY 86 YORSQPIWGLHDPKQKQ-----WQWIDGAMYLYRWSG--KSMGNGKHCAEWSNNNF 138  
DB 95 VNNQDIWIWLHDPMTGQQPNGGGWWSNDVLNLYNWDGDPSTVNRGNCGLTATSEF 154  
QY 139 LTWSNNECNKROHFLCKYR 157  
DB 155 LKWDGHHCDVLEPFVCKFK 173  
RESULT 11  
RGHULA  
regenerating islet lectin 1-alpha precursor [validated] - human  
N;Alternate names: lithostathine; pancreatic thread protein (PTP); reg I protein; reg1-a1  
N;Contains: pancreatic stone protein (PSP)  
C;Species: Homo sapiens (man)  
C;Date: 31-Mar-1990 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004  
C;Accession: A35197; B28351; S12950; S02767; S01413; S01471; A25246  
R;Watanabe, T.; Yonekura, H.; Terazono, K.; Yamamoto, H.; Okamoto, H.  
J. Biol. Chem. 265, 7432-7439, 1990  
A;Title: Complete nucleotide sequence of human reg gene and its expression in normal and product of the gene.  
A;Reference number: A35197; MUID:90237042; PMID:2332435  
A;Accession: A35197  
A;Molecule type: DNA  
A;Residues: 1-166 <WAT>  
A;Cross-references: UNIPROT:P05451; GB:J05412  
R;Terazono, K.; Yamamoto, H.; Takagawa, S.; Shiga, K.; Yonemura, Y.; Tochino, Y.; Okamoto, K.  
J. Biol. Chem. 265, 2111-2114, 1990  
A;Title: A novel gene activated in regenerating islets.  
A;Reference number: A92704; MUID:88115343; PMID:2963000  
A;Accession: B28351  
A;Molecule type: mRNA  
A;Residues: 1-166 <TER>  
A;Cross-references: GB:M18963; NID:g190978; PIDN:AAA36558.1; PID:g190979  
R;Itoh, T.; Tsuzuki, H.; Katoh, T.; Teraoka, H.; Matsumoto, K.; Yoshida, N.; Terazono, K.  
FEBS Lett. 272, 85-88, 1990  
A;Title: Isolation and characterization of human reg protein produced in Saccharomyces cerevisiae.  
A;Reference number: S12950; MUID:91032149; PMID:2226837  
A;Accession: S12950  
A;Molecule type: protein  
A;Residues: 23-52;160-166 <ITO>  
A;Note: sequence determined from protein isolated after human cDNA sequence was cloned and expressed in E. coli.  
R;de Caro, A.M.; Adrich, Z.; Fournet, B.; Capon, C.; Bonicel, J.J.; de Caro, J.D.; Roverly, J.  
Biochim. Biophys. Acta 994, 281-284, 1989  
A;Title: N-terminal sequence extension in the glycosylated forms of human pancreatic stor protein.  
A;Reference number: S02767; MUID:89150292; PMID:2493268  
A;Accession: S02767  
A;Molecule type: protein  
A;Residues: 23-47 <DEC>  
R;Rouimi, P.; de Caro, J.; Bonicel, J.; Roverly, M.; de Caro, A.



A;Accession: A45751  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-166 <GIO>  
A;Cross-references: UNIPROT:P05451; GB:M27190; NID:G623412; PIDN:AAA60546.1; PID:G623413  
C;Superfamily: tetranectin; C-type lectin homology  
F;36-162/Domain: C-type lectin homology <LCH>

Query Match 25.5%; Score 224; DB 2; Length 166;  
Best Local Similarity 33.3%; Pred. No. 9.4e-15;  
Matches 45; Conservative 26; Mismatches 58; Indels 6; Gaps 4;

QY 27 RPSCAPGWYHKSNCYGYFKLRNWSDALEQCQSYGNGAHLASILSLKEASTIAEVI--S 84  
DB 33 RISCPEGTWAYRSYCYFFNEDETWDADLYQNNWSG-NLVSIVLTQAEAFVASLIKES 91  
QY 85 GYRSQPIWIGLHDPKQRQOWIDGAMLYLRSW--SGKSMGNGKHCAMSSNNNNFLTWS 142  
DB 92 GTDDFN-VVIGLHDPKKNRHWSSGSLVYSKSWGIGAPSSVNPQVCVSLTSSSTGFQKWK 150  
QY 143 SNECKNRQHFCLKYR 157  
DB 151 DVPCEDKFSFVCKFK 165

RESULT 13  
S32489 lectin - Iberian ribbed newt  
C;Species: Pleurodeles waltlii (Iberian ribbed newt)  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S32489; S28530  
R;Tiffiche, C.; Chesnel, A.; Jegou, J. P.  
Eur. J. Biochem. 213, 901-907, 1993  
A;Title: Isolation and characterization of a cDNA clone encoding a Pleurodeles lectin.  
A;Reference number: S32489; MUID:93279340; PMID:8504829  
A;Accession: S32489  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-172 <TIF>  
A;Cross-references: UNIPROT:Q02988; EMBL:X69062; NID:G64257; PIDN:CAA48800.1; PID:G64258  
C;Superfamily: tetranectin; C-type lectin homology  
F;37-164/Domain: C-type lectin homology <LCH>

Query Match 25.3%; Score 222.5; DB 2; Length 172;  
Best Local Similarity 34.3%; Pred. No. 1.4e-14;  
Matches 45; Conservative 26; Mismatches 58; Indels 6; Gaps 4;

[illegible]

pancreatitis-associated protein precursor - human  
N/Alternate names: C-type lectin; pancreatic stone protein homolog HIP  
C/Species: Homo sapiens (man)  
C/Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text\_change 09-Jul-2004  
C/Accession: A49616; A44931; S29821; S48197; I55580  
R/Dusetti, N.J.; Frigerio, J.M.; Fox, M.F.; Swallow, D.M.; Dagnorn, J.C.; Iovanna, J.L.  
Genomics 19, 108-114, 1994  
A/Title: Molecular cloning.  
A/Reference number: A49616; MUID:94245143; PMID:8188210  
A/Accession: A49616  
A/Molecule type: DNA

A;Residues: 1-175 <DUS>  
A;Cross-references: UNIPROT:Q06141; GB:LI15533; NID:G482908; PIDN:AAA60020.1; PID:G482909  
R;LaSalette, C.; Christa, L.; Simon, M.T.; Vernier, P.; Brechot, C.  
Cancer Res. 52, 5089-5095, 1992  
A;Title: A novel gene (HIP) activated in human primary liver cancer.  
A;Reference number: A44931; MUID:92386513; PMID:1325291  
A;Accession: A44931  
A;Molecule type: mRNA  
A;Residues: 1-175 <LAS>  
A;Cross-references: GB:X68641; NID:G312806; PIDN:CAA48605.1; PID:G312807  
A;Experimental source: hepatocellular carcinoma  
A;Note: sequence extracted from NCBI backbone (NCBIN:113007, NCBIP:113008)  
R;Itoh, T.; Teraoka, H.  
Biochim. Biophys. Acta 1172, 184-186, 1993  
A;Title: Cloning and tissue-specific expression of cDNAs for the human and mouse homolog  
A;Reference number: S29821; MUID:93176807; PMID:767928  
A;Accession: S29821  
A;Status: Preliminary  
A;Molecule type: mRNA  
A;Residues: 1-175 <TO>  
A;Cross-references: GB:DL13510; NID:G285970; PIDN:BA02728.1; PID:G285971  
R;LaSalette, C.; Simon, M.T.; Ishikawa, H.; Dirlong, S.; Nguyen, V.C.; Christa, L.; Vernier, P.  
Eur. J. Biochem. 224, 29-38, 1994  
A;Title: Structural organization and chromosomal localization of a human gene (HIP/PAP)  
A;Reference number: S48197; MUID:94357229; PMID:8076648  
A;Accession: S48197  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-175 <LA2>  
R;Orell, B.; Keim, V.; Masciotra, L.; Dagorn, J.C.; Iovanna, J.L.  
J. Clin. Invest. 90, 2284-2291, 1992  
A;Title: Human pancreatitis-associated protein. Messenger RNA cloning and expression in  
A;Reference number: I55580; MUID:93107309; PMID:1469087  
A;Accession: I55580  
A;Status: Preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-175 <RES>  
A;Cross-references: GB:S51768; NID:G262368; PIDN:AA824642.1; PID:G262369  
C;Genetics:  
A;Gene: GDB:PAP; HIP  
A;Cross-references: GDB:136839; OMIM:167805  
A;Map position: 2p12-2p12  
A;Introns: 26/1; 65/3; 111/3; 154/1  
C;Superfamily: tetranectin; C-type lectin homology  
F;Keywords: acute phase; extracellular protein; pancreas  
F;1-26/Domain: signal sequence #status predicted <SIG>  
F;27-175/Product: pancreatitis-associated protein #status predicted <MAT>  
F;40-171/Domain: C-type lectin homology <LCH>  
F;40-51,68-171,146-163/Disulfide bonds: #status predicted

Query Match 25.2%; Score 221; DB 2; Length 175;  
Best Local Similarity 32.7%; Pred. No. 2e-14;  
Matches 53; Conservative 29; Mismatches 64; Indels 16; Gaps 7;

QY 10 LLLSCL-ARTGVLGDIIMR-PS- ---CAPGWFYHKSNCYGFRKLNWSDAELECSYVG 62  
Db 13 MLLSCLMLLSQVQBEERQRLPRLRCPKGSAKXGSHYALFLSPKSWTDADLACQKRP 72

QY 63 NGHILASLTLKEASTIAEYISGYQSQP- IWIGLHDPQKQOQ-----QWIDGMYLYRWSGKSNMG 116  
Db 73 SG-NLVSVLSGAEQSFVSSLVKSGISYGYWIGLHDPQKQOQ-----QWIDGMYLYR 131

QY 117 SW-SGKSMGNGKICAEKMSNNPLTWSSNECNKQHFPLCKY 156  
Db 132 AWERNPSTISSFGHCASLSRSTAFRLRWKDYNCNVLFPYCKF 173

RESULT 15  
JC7134  
agkisacutacin alpha chain precursor - sharp-nosed viper  
N;Alternate names: fibrinogenolytic venom protein  
C;Species: Agkistrodon acutus (sharp-nosed viper)  
C;Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 09-Jul-2004

C;Accession: JC7134; PC7037  
R;Cheng, X.; Qian, Y.; Liu, Q.; Li, B.X.Y.; Zhang, M.; Liu, J.  
Biochem. Biophys. Res. Commun. 265, 530-535, 1999  
A;Title: Purification, characterization, and cDNA cloning of a new fibrinogenolytic venom  
A;Reference number: JC7134; MUID:20025379; PMID:10558903  
A;Accession: JC7134  
A;Molecule type: mRNA  
A;Residues: 1-152 <CHE>  
A;Cross-references: UNIPROT:Q9DEF9; UNIPROT:Q8JIW0; GB:AF176420  
A;Experimental source: venom gland  
A;Accession: PC7037  
A;Molecule type: protein  
A;Residues: 24-53;84-86;87-94;125-136;137-152 <CH2>  
C;Superfamily: tetranectin; C-type lectin homology  
F;Keywords: disulfide bond; heterodimer; venom  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-152/Product: agkisacutacin alpha chain #status experimental <MAT>

Query Match 23.5%; Score 206.5; DB 2; Length 152;  
Best Local Similarity 31.3%; Pred. No. 4.5e-13;  
Matches 47; Conservative 26; Mismatches 64; Indels 13; Gaps 4;

QY 10 LLLSCLARTGVLGDIIMRPSCAPGWFYHKSNCYGFRKLNWSDAELECSYGNCAHLAS 69  
Db 11 LLVVFLSLSGTAAD-----CSSGWSVYEGHCYKVFQSKTWADAESFCTKQVNGHLVS 64

QY 70 ILSLKEASTIAEYISGYQSQPI--WIGLHDPQKQOQ--QWIDGMYLYRWSGKSNMG 125  
Db 65 IESSGEADFVAHLIAQKIKSAKIHVIGLRAQNKKEKQCSIEWSDGSSISYENWIEE--- 121

QY 126 NKHCAEMSSNNPLTWSSNECNKQHFPLCK 155  
Db 122 SKKCLGVHETGFKWENFYCEQQDPFVCE 151

Search completed: March 3, 2005, 07:55:08  
Job time : 40 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 07:50:03 ; Search time 177 Seconds  
(without alignments)  
457.110 Million cell updates/sec

Title: US-10-099-791E-2

Perfect score: 878

Sequence: 1 MASRSMRLLLLSCLAKTGV.....LTWSSNECNKRQHFCLKYRP 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878	100.0	158	2 Q9BYZ8	Q9BYZ8 homo sapien
2	600.5	68.4	157	2 Q9D805	Q9D805 mus musculus
3	594.5	67.7	157	2 Q9D858	Q9D858 mus musculus
4	593.5	67.6	157	2 Q68AX7	Q68AX7 rattus norv
5	544	62.0	113	2 Q8NER7	Q8NER7 homo sapien
6	438.5	49.9	160	2 Q7SZ75	Q7SZ75 xenopus lae
7	307.5	35.0	134	2 Q8NER6	Q8NER6 homo sapien
8	260.5	29.7	132	1 STR1_STRCA	P83514 struthio ca
9	258.5	29.4	132	1 ACAL_ANGAN	P83300 anser anser
10	254.5	29.0	165	1 LIT1_MOUSE	P43137 mus musculus
11	254.5	29.0	165	1 LIT1_MOUSE	P43137 mus musculus
12	247	28.1	142	1 STR2_STRCA	P83515 struthio ca
13	246.5	28.1	174	1 PAP3_MOUSE	Q09049 mus musculus
14	245	27.9	164	2 Q6TR36	Q6TR36 bothrops ja
15	243.5	27.7	158	2 Q6QX33	Q6QX33 bothrops in
16	242	27.6	175	1 PAP2_MOUSE	Q09037 mus musculus
17	241.5	27.5	142	1 OC17_CHICK	Q3PR88 gallus gall
18	240	27.3	166	1 LITB_HUMAN	P48304 homo sapien
19	237.5	27.1	158	2 Q7IRQ1	Q7IRQ1 trimeresuru
20	237	27.0	135	1 LECG_BOTJR	P83519 bothrops ja
21	236.5	26.9	173	1 LIT2_MOUSE	Q08731 mus musculus
22	236	26.9	126	2 Q7C278	Q7C278 bothrops ja
23	235	26.8	126	2 Q8C6F9	Q8C6F9 mus musculus
24	234	26.7	135	1 LECG_BITAR	Q9P890 bitis ariet
25	233	26.5	135	1 LECG_CROAT	P21963 crotalus at
26	231	26.3	174	1 PAP3_RAT	P42854 rattus norv
27	231	26.3	175	1 LITB_BOVIN	P23132 bos taurus
28	230.5	26.3	154	2 Q8JIV9	Q8JIV9 agkistrodon
29	229	26.1	135	1 LECG_LACST	Q9P8M4 lachesis st
30	228.5	26.0	174	1 PAP2_RAT	P35231 rattus norv
31	227.5	25.9	146	2 Q9CVF4	Q9CVF4 mus musculus

32 227 25.9 158 2 Q90WI7 Q90WI7 bungarus fa  
33 227 25.9 158 2 Q90WI8 Q90WI8 bungarus fa  
34 225.5 25.7 146 2 Q6X5S1 Q6X5S1 echis pyram  
35 225.5 25.7 146 2 Q6X5S4 Q6X5S4 echis carin  
36 225.5 25.7 166 1 LITA\_HUMAN P05451 homo sapien  
37 222.5 25.3 172 1 LECG\_PLEWA Q02988 pleurodeles  
38 222 25.3 157 2 Q6T7B5 Q6T7B5 bitis gabon  
39 221.5 25.2 155 2 Q8JIV8 Q8JIV8 agkistrodon  
40 221 25.2 175 1 PAP1\_HUMAN Q06141 homo sapien  
41 219.5 25.0 148 2 Q6X5S3 Q6X5S3 echis pyram  
42 219.5 25.0 148 2 Q6X5S7 Q6X5S7 echis ocell  
43 219.5 25.0 148 2 Q6X5S9 Q6X5S9 echis carin  
44 219.5 25.0 148 2 Q6X5T1 Q6X5T1 bitis ariet  
45 218.5 24.9 154 2 Q7T2Q0 Q7T2Q0 echis multi

## ALIGNMENTS

### RESULT 1

Q9BYZ8 PRELIMINARY; PRT; 158 AA.  
AC Q9BYZ8:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
DE Regenerating gene type IV precursor (REG-like protein) (Regenerating  
DE islet-derived family, member 4) (Gastrointestinal secretory protein  
DE GISP).  
GN Name=REG4;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21210973; PubMed=11311942; DOI=10.1016/S0167-4781(00)00284-0;  
RA Hartup J.C., Zhang H., Bonaldo M.F., Soares M.B., Dieckgraefe B.K.;  
RT "Isolation and characterization of a cDNA encoding a novel member of  
RT the human regenerating protein family: Reg IV(1).";  
RL Biochim. Biophys. Acta 1518:287-293(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon cancer;  
RA Violette S., Festor E., Dussaulx E., Citadelle D., Chambaz J.,  
RA Lacasa M., Lesuffleur T.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Kamarainen M., Heiskala K., Heiskala M., Andersson L.C.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,  
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Heltan E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human



```

[8]
RN  SEQUENCE FROM N.A.
RP  STRAIN=FVB/N; TISSUE=Colon;
RA  Strauberg R.;
RL  Submitted (DfC-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AK008049; BAB25429.1; -
DR  EMBL; BC019465; ANH19465.1; -
DR  HSSP; Q06141; IUV0.
DR  MGD; MGI:1914959; Reg4.
DR  GO; GO:0005615; C:extracellular space; TAS.
DR  InterPro; IPR001304; Lectin_C.
DR  Pfam; PF00059; Lectin_C_1.
DR  SMART; SM00034; CLECT; 1.
DR  PROSITE; PS00041; C TYPE LECTIN 2; 1.
SQ  SEQUENCE 157 AA; 18398 MW; F3981722BBD83968 CRC64;

Query Match 68.4%; Score 600.5; DB 2; Length 157;
Best Local Similarity 66.2%; Pred. No. 4.8e-51;
Matches 104; Conservative 25; Mismatches 27; Indels 1; Gaps 1;

Qy 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCGYFRKLRNWSDAELSCQS 60
Db 1 MASGVRLLLLLSWAGPEVLSD-ILRPSCAPGWFYRSHCYGFRKLRNWSHAELSCQS 59

Qy 61 YNGAHLASLILSLKEASTIAEYISGYQSORPIWGLHDPOKQOWIDGAMYLRSWSG 120
Db YGNGSHLASVLNQKEASVISKYITGYQRLNPLVWGLHDPOKQQLQWTDGSTNLYRRNP 119

Qy 121 KSMGNGKHCAEMSNNNFLTWSSNECKRQHFLCKYR 157
Db 120 RTKSEARHCAEMNPDKRFLTNKNGCANRQHFLCKYK 156

RESULT 3
Q9D858 PRELIMINARY; PRT; 157 AA.
AC Q9D858;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male small intestine cDNA, RIKEN full-length
DE enriched library, clone:2010204K21 product:REGENERATING GENE TYPE IV,
DE full insert sequence.
GN Name=Reg4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=C57BL/6J; TISSUE=Small intestine;
RX  MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA  Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
[2]
RN  SEQUENCE FROM N.A.
RP  STRAIN=C57BL/6J; TISSUE=Small intestine;
RX  MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA  RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
[3]
RN  SEQUENCE FROM N.A.
RP  STRAIN=C57BL/6J; TISSUE=Small intestine;
RA  The FANTOM Consortium,
RA  the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
[4]

```

```

RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Small intestine;
RX  MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA  Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA  Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
[5]
RN  SEQUENCE FROM N.A.
RP  STRAIN=C57BL/6J; TISSUE=Small intestine;
RX  MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA  Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA  Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA  Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA  Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA  Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Matakaki M.,
RA  Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA  Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
[6]
RN  SEQUENCE FROM N.A.
RP  STRAIN=C57BL/6J; TISSUE=Small intestine;
RC  Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA  Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA  Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA  Inotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA  Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA  Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA  Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA  Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA  Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA  Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA  Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AK008438; BAB25669.1; -
DR  HSSP; Q06141; IUV0.
DR  MGD; MGI:1914959; Reg4.
DR  GO; GO:0005615; C:extracellular space; TAS.
DR  InterPro; IPR001304; Lectin_C.
DR  InterPro; IPR003990; Pancreatins_ac.
DR  Pfam; PF00059; Lectin_C_1.
DR  PRINTS; PR01504; PNCREATITSAP.
DR  SMART; SM00034; CLECT; 1.
DR  PROSITE; PS00041; C TYPE LECTIN 2; 1.
SQ  SEQUENCE 157 AA; 18474 MW; FD96F36CFB989368 CRC64;

Query Match 67.7%; Score 594.5; DB 2; Length 157;
Best Local Similarity 65.6%; Pred. No. 1.9e-50;
Matches 103; Conservative 25; Mismatches 28; Indels 1; Gaps 1;

Qy 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCGYFRKLRNWSDAELSCQS 60
Db 1 MAYKGVRLLLLLSWAGPEVLSD-ILRPSCAPGWFYRSHCYGFRKLRNWSHAELSCQS 59

Qy 61 YNGAHLASLILSLKEASTIAEYISGYQSORPIWGLHDPOKQOWIDGAMYLRSWSG 120
Db YGNGSHLASVLNQKEASVISKYITGYQRLNPLVWGLHDPOKQQLQWTDGSTNLYRRNP 119

Qy 121 KSMGNGKHCAEMSNNNFLTWSSNECKRQHFLCKYR 157
Db 120 RTKSEARHCAEMNPDKRFLTNKNGCANRQHFLCKYK 156

RESULT 4
Q68AX7 PRELIMINARY; PRT; 157 AA.
AC Q68AX7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

```

```

DE GN Name=Reg4;
OS Rattus norvegicus (Rat);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=ileum;
RA Namikawa K., Murakami K., Fukushima M., Kiyama H.;
RT "Differential regulation of Reg family member expression after
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB164049; BAD38673.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF000390; Pancreatins_ac.
DR PRINTS; PR01504; PNCREATINSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C-TYPE LECTIN 2; 1.
DR SEQUENCE 157 AA; 18369 MW; 55B129FB10BA4D1D CRC64;

Query Match 67.6%; Score 593.5; DB 2; Length 157;
Best Local Similarity 66.5%; Pred. No. 2.3e-50;
Matches 105; Conservative 21; Mismatches 31; Indels 1; Gaps 1;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGFWYHKSNCYGYFRKLRNWSDAELECOS 60
Db 1 MASKCVRLLLLSWVAGPEVLSD-ILRPSGASGFWYHSHCYGYFRKLRNWSHAELECOS 59
||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 YONGAHLASILSLKEASTIAEYISGYORSOPITWIGLHDPKQR 120
YONGAHLASVLNPKASVSKITYTQYRSLPVMVIGLHDPQKASQWIDGSTNQYRPNSP 119
||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 YNGSGHLASVLNPKASVSKITYTQYRSLPVMVIGLHDPQKASQWIDGSTNQYRPNSP 119
||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 KSMGNGKHCAMSNNNFLTWSSNECNKQHFCLKYRP 158
120 RTKSBARHCETMNPDKRFLTWNNKNGCTKQHFCLKYRP 157
:::|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
Q8NER7 PRELIMINARY; PRT; 113 AA.
AC Q8NER7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE REG-like protein splice variant 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamarainen M., Heiskala K., Heiskala M., Andersson L.C.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY126671; AAM95599.1; -.
DR HSP; P21963; 1J2N.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF000390; Pancreatins_ac.
DR PRINTS; PR01504; PNCREATINSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C-TYPE LECTIN 2; 1.
DR SEQUENCE 113 AA; 12832 MW; AZSDFFIA729C78DA CRC64;

Query Match 62.0%; Score 544; DB 2; Length 113;
Best Local Similarity 99.0%; Pred. No. 1.2e-45;
Matches 101; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGFWYHKSNCYGYFRKLRNWSDAELECOS 60
Db 1 MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGFWYHKSNCYGYFRKLRNWSDAELECOS 60
||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 YONGAHLASILSLKEASTIAEYISGYORSOPITWIGLHDPKQR 102
YONGAHLASVLNPKASVSKITYTQYRSLPVMVIGLHDPQKASQWIDGSTNQYRPNSP 102
||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 YONGAHLASILSLKEASTIAEYISGYORSOPITWIGLHDPKQR 102
YONGAHLASVLNPKASVSKITYTQYRSLPVMVIGLHDPQKASQWIDGSTNQYRPNSP 102
||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
Q7SZ75 PRELIMINARY; PRT; 160 AA.
AC Q7SZ75;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MGC64513 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Kautner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Rickard S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053817; AAHS3817.1; -.
DR HSP; P22897; IEGG.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_C; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00645; C-TYPE LECTIN 1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN 2; 1.
DR SEQUENCE 160 AA; 18330 MW; 6A5502F24689179A CRC64;

Query Match 49.9%; Score 438.5; DB 2; Length 160;
Best Local Similarity 48.7%; Pred. No. 4.1e-35;
Matches 75; Conservative 35; Mismatches 41; Indels 3; Gaps 2;

QY 8 ILLLSCLAKTGVLDIIMRPSCAPGFWYHKSNCYGYFRKLRNWSDAELECOSGNAHL 67
||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 7 LLLPLGALVSNVLEAAQVRSSCPNGHFFKANCYGYFRYPLSWAEYDCQAYGHGAHL 66
Qy 68 ASILSLKEASTIAYISGYQSRQPIWGLHDPKQKQOWIDGMYLYRSW-SGKSMGNG 126
Db 67 ASILSDAEADVIASHAYQKNFVWIGLHDPEQNRKRWKNDGSMYRSLWLAGQPNYN 126
Qy 127 --KHCAEMSNNNFLTWSSNECKRQHFLCKYRP 158
Db 127 SABYCGELCKEGFVKRWNSCKEYKQYVCKYRP 160

RESULT 7
Q8NER6
ID Q8NER6 PRELIMINARY; PRT; 134 AA.
AC Q8NER6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE REG-like protein splice variant 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamarainen M., Heiskala K., Heiskala M., Andersson L.C.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY126672; AAM95600.1;
SQ SEQUENCE 134 AA; 14993 MW; E0E5AD9B96A53EBS CRC64;

Query Match 35.08; Score 307.5; DB 2; Length 134;
Best Local Similarity 55.94; Pred. No. 2.6e-22;
Matches 62; Conservative 9; Mismatches 9; Indels 31; Gaps 3;

Qy 1 MASRMLLLLLSLCTAGTGLDIIIMPSCAPGWFYHKSNCYGYFRKLRNWSDAEL--- 56
Db 1 MASRMLLLLLSLCTAGTGLDIIIMPSCAPGWFYHKSNCYGYFRKLRNWSDAEVRNLL 60
Qy 57 -----ECQSYGNGAHLASILSLKEASTIAYISGYQSRQPIWI 94
Db 61 PAWPLSLRAXDQPEQP-----ISFDSGSSV---LPGHYEKLPL 97

RESULT 8
STR1_STRCA
ID STR1_STRCA STANDARD; PRT; 132 AA.
AC P8351a;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DE Struthioalcin-1 (SCA-1).
OS Struthio camelus (Ostrich).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;
OC Struthio.
OX NCBI_TaxID=8801;
RN [1]
RP SEQUENCE, SUBCELLULAR LOCATION, AND MASS SPECTROMETRY.
RC TISSUE=Eggshell matrix;
RX PubMed=14726203; DOI=10.1016/j.bbapap.2003.09.006;
RA Mann K., Siedler F.;
RT "Ostrich (Struthio camelus) eggshell matrix contains two different C-
type lectin-like proteins. Isolation, amino acid sequence, and
posttranslational modifications."
RL Biochim. Biophys. Acta 1696:41-50(2004).
CC -1- SUBCELLULAR LOCATION: Eggshell matrix.
CC -1- MASS SPECTROMETRY: MW=15343.2; MW_ERR=4; METHOD=Electrospray;
CC RANGE=1-132; NOTE=Ref.1.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR001304; Lectin_C.
KW Direct protein sequencing; Lectin.
FT DOMAIN 1 132 C-type lectin.
FT DISULFID 3 14 By similarity.

```

```

DR PRINTS; PRO0356; ANTIFREEZEII.
DR PRINTS; PRO1504; PNCREATISAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C TYPE LECTIN 1; 1.
DR PROSITE; PS00041; C TYPE LECTIN 2; 1.
KW Direct protein sequencing; Lectin.
FT DOMAIN 10 129 C-type lectin.
FT DISULFID 3 14 By similarity.
FT DISULFID 31 128 By similarity.
FT DISULFID 103 120 By similarity.
SQ SEQUENCE 132 AA; 15353 MW; F7BD1DF2990B2945 CRC64;

Query Match 29.78; Score 260.5; DB 1; Length 132;
Best Local Similarity 38.88; Pred. No. 1.1e-17;
Matches 50; Conservative 20; Mismatches 54; Indels 5; Gaps 2;

Qy 30 CAPGWFYHKSNCYGYFRKLRNWSDAELQSCYNGAHLASILSLKEASTIAYISGY--- 86
Db 3 CPKGLDFRCGNCYGYFRYELPWKRAEAWCRSIRAGAHLASIHTSEHRAIAKPIQYHHG 62
Qy 87 QRSQPIWGLHDPKQKQOWIDGMYLYRSWKGKSMGNGKHCMAEMSNNNFLTWSSNEC 146
Db 63 EEBEDVWIGLF--RWNSVMAWIDGSKKHSALDDDDYPKGKHCVALDESSGFLSDNDSC 120
Qy 147 NKQHFELCK 155
Db 121 GERNAFICK 129

RESULT 9
ACAL_ANSAN
ID ACAL_ANSAN STANDARD; PRT; 132 AA.
AC P83300;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ansoalcin.
OS Anser anser anser (Western graylag goose).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.
OX NCBI_TaxID=8844;
RN [1]
RP SEQUENCE, FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, AND MASS
SPECTROMETRY.
RC TISSUE=Eggshell matrix;
RX MEDLINE=22439773; PubMed=12431998; DOI=10.1074/jbc.M201518200;
RA Lakshminarayana R., Vallyaveetil S., Rao V.S., Kini R.M.;
RT "Purification, characterization, and in vitro mineralization studies
of a novel goose eggshell matrix protein, ansoalcin."
RL J. Biol. Chem. 278:2928-2936(2003).
CC -1- FUNCTION: Induces spherical aggregates of calcite crystals in
vitro. Believed to play an active role in the eggshell
calcification.
CC -1- SUBUNIT: Homodimer or homotrimer.
CC -1- SUBCELLULAR LOCATION: Eggshell matrix. May be present in the shell
glands on the walls of oviduct and incorporated into the shell
structure during its formation.
CC -1- MASS SPECTROMETRY: MW=15342; METHOD=Electrospray; RANGE=1-132;
CC NOTE=Ref.1.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
DR HSP; P23806; IJ34.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; Lectin_C; 1.
DR PRINTS; PRO0356; ANTIFREEZEII.
DR PRINTS; PRO1504; PNCREATISAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C TYPE LECTIN 1; 1.
DR PROSITE; PS00041; C TYPE LECTIN 2; 1.
KW Direct protein sequencing; Lectin.
FT DOMAIN 1 132 C-type lectin.
FT DISULFID 3 14 By similarity.

```









```

Db 123 ATWDELCSDRKPFICBYR 141

RESULT 13
PAP3_MOUSE
ID PAP3_MOUSE STANDARD; PRT; 174 AA.
AC Q09049;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pancreatitis-associated protein 3 precursor (REG III-gamma).
GN Name-Pap3; Synonyms=Reg3g;
OS Mus musculus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver, and Pancreas;
RX MEDLINE=97208868; PubMed=9055810; DOI=10.1016/S0378-1119(96)00589-6;
RA Narushima Y., Unno M., Nakagawa K.-I., Mori M., Miyashita H.,
RA Suzuki Y., Noguchi N., Takasawa S., Kumagai T., Yonekura H.,
RA Okamoto H.;
RT "Structure, chromosomal localization and expression of mouse genes
encoding type III Reg, RegIII alpha, RegIII beta, RegIII gamma.";
RL Gene 185:159-168(1997).
CC -!- FUNCTION: Might be a stress protein involved in the control of
bacterial proliferation.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Constitutively expressed in the small
intestine, moderately in colon and at an extremely low level in
healthy pancreas.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; D63361; BAA18930.1; -.
CC DR EMBL; D63362; BAA18931.1; -.
CC DR HSSP; P05451; 111T.
CC DR MGD; MGI:109406; Reg3g.
CC DR InterPro; IPR001304; Lectin C.
CC DR Pfam; PF00059; Lectin C; 1.
CC DR PRINTS; PR01504; PNCREATITSAP.
CC DR SMART; SM00034; CLECT.1.
CC DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
CC DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
CC FT CHAIN 24 157 BJCUL.
CC SQ SEQUENCE 164 AA; 19070 MW; F88101D4338B94DC CRC64;

Query Match 27.9%; Score 245; DB 2; Length 164;
Best Local Similarity 36.9%; Pred. No. 4.5e-16;
Matches 48; Conservative 16; Mismatches 62; Indels 4; Gaps 2;

QY 29 SCAPGFYFKSNICYGFRKLRNWSDAELECOSYNGNGAHLASILSLKEASTIAYISGYOR 88
Db 25 NCPQDWLPNGLCYKIFNELKAWKDAEMFCRKYKPGCHLASIHLGYESPEIAYISDYHK 84
QY 89 SQ-PIWIGLHDPKROQWOWIDGAMLYRSWSGKS---MGKNKHCAMSSNNFLTWSN 144
Db 85 GQSEVMIGLCDDKKDFSEWTDRTSCTDYLSDWKNQPDHYQNKFCVELVSNTRYRLMNDQ 144
QY 145 ECKRQHFLC 154
Db 145 VCESKNFLC 154

RESULT 15
Q6QX33
ID Q6QX33 PRELIMINARY; PRT; 158 AA.
AC Q6QX33;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE C-type lectin.
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=venom gland;

Query Match 28.1%; Score 245.5; DB 1; Length 174;
Best Local Similarity 35.2%; Pred. No. 3.4e-16;
Matches 57; Conservative 24; Mismatches 66; Indels 15; Gaps 6;

QY 10 LLLSCL-AKTGVLDII-----MRPSCAPGFYFKSNICYGFRKLRNWSDAELECOSYG 62
Db 13 MLLSCLMLLSQVGEVAKDAPSRSSCPKGRAYGVCYALFVSKVKNYADADMACQKRP 72
QY 63 NGAHLASTLSLKEASTIAYI-SGYQROPIWIGLHDP-----QKROQWOWIDGAMLYR 116
Db 73 SG-HLVSVLSGAEASFLSSMKSGNSGQYVWIGLHDPITLGYEENRGWENSNADVWYI 131

```

Search completed: March 3, 2005, 07:53:08  
Job time : 179 secs

